Understanding, normativity, and scientific practice

LEWENDON-EVANS, HARRY,EDWARD

Understanding, Normativity, and Scientific Practice

Harry Lewendon-Evans

PhD Thesis
Department of Philosophy
Durham University
2018

Abstract

Understanding, Normativity, and Scientific Practice

Harry Lewendon-Evans

Recent work in epistemology and philosophy of science has argued that understanding is an important cognitive achievement that philosophers should seek to address for its own sake. This thesis outlines and defends a new account of scientific understanding that analyses the concept of understanding in terms of the concept of normativity. The central claim is that to understand means to grasp something in the light of norms. The thesis is divided into two parts: Part I (chapters one to three) addresses the question of the agency of understanding and Part II (chapters four to five) focuses on the vehicles of scientific understanding. Chapter One begins with an account of understanding drawn from the work of Martin Heidegger, which presents understanding as a practical, normative capacity for making sense of entities. Chapter Two builds on Robert Brandom's normative inferentialism to argue that conceptual understanding is grounded in inferential rules embedded within norm-governed, social practices. Chapter Three argues that normativity should be located in the intersubjective nature of social practices. The chapters in Part II draw on and extend the account of understanding developed in Part I by focusing on how models and explanations function within scientific practice to facilitate scientific understanding. Chapter Four investigates the nature of model-based understanding. It defends the claim that constructing and using models enables a form of conceptual articulation which facilitates scientific understanding by rendering scientific phenomena intelligible. Chapter Five considers the connection between understanding and explanation through the role of explanatory discourse in scientific practice. I argue that the function of explanations is to sculpt and make explicit the norms of intelligibility required for scientific understanding. This thesis concludes that scientific understanding is an inherently norm-governed phenomenon that is unintelligible without reference to the normative dimension of our social and scientific practices.

Table of Contents

Acknowledgements

First of all, I would like to thank my supervisors Simon James and Peter Vickers for their unwavering support, patience and guidance over the last four and a half years. Both have provided me with the confidence to pursue the arguments developed in this thesis, as well as the space and encouragement to develop them in all the weird and wonderful directions that make philosophy worth doing in the first place. I'd also like to thank the Arts and Humanities Research Council for generously funding this thesis.

The Philosophy Department in Durham has provided a fantastic environment to work, learn, and teach in. In particular, I'd like to thank Ben Smith for his support in putting together the proposal for the PhD as well as for many helpful conversations over the years. I am indebted to Ian Kidd for his constant encouragement and for the long, free-wheeling chats over coffee. I'm extremely grateful to Wendy Parker for reading and commenting on drafts of the scientific models chapter, which helped to tighten up the arguments. I'd also like to thank Andrew Cooper, Matthew Eddy, Matthew Tugby, and Richard Stopford for their advice, both on an academic and professional level. Finally, I want to thank Andy Hamilton for the jazz jams.

Philosophy is fundamentally social; ideas and arguments rarely develop and flourish on their own. Instead, they are built and enriched through interaction and collaboration and by being challenged and encouraged by others. In this respect, this thesis owes much to the department's postgraduate community. This thesis originally began life as a work on Heidegger and the development of his thoughts on science, but through conversations with other postgraduates working on philosophy of science, my interests significantly expanded and the focus shifted. I would therefore like to thank all those over the years who have had a formative influence on the thesis: Josh Bergamin, Anna Bortolan Anna de Bruyckere, Anthony Fernandez, Holly Havens, Salomé Jacobs, William Jones, Lee Large, Erin Nash, Rune Nyrup, William Peden, Tom Rossetter, and Sarah Wieten. I would also like to thank audiences in Durham, London, Leeds, Nottingham, Atlanta, Copenhagen, Parnu for insightful questions and discussion, and in particular, Michael T Stuart for his extremely helpful comments on an earlier version of the chapter on explanation and understanding.

The broader life and community of Durham is focused through its colleges, and in that regard I would like to thank the community of St Cuthbert's Society, and in particular Elizabeth Archibald, Sharon Richardson, Phil Bolton, Andrew Tibbs, and Jon Warren for their wonderful support and encouragement during the PhD. I am also very grateful to Cuth's for providing me with employment when funding ran out and the flexibility to continue to work on the thesis.

I would also like to thank all the friends who have listened to me rattle on about understanding, norms, scientific practices, models, and phenomenology during the last few years and offered much needed distraction. In particular, I would like to thank Flora Hetherington, who introduced me to some fascinating issues in plant science and offered a scientist's perspective on my arguments and to John Foxwell, who not only discussed and challenged some of my most basic ideas and assumptions, but who also read and commented on various chapter drafts.

I'd especially like to thank my family, who have been brilliant and wonderfully supportive throughout the PhD. My parents, Julie and Keith, have been a constant source of unconditional love and inspiration and I can only thank them for all they have done for me. Finally, I'd like to thank Lara. Without you this thesis would never have been written. I dedicate it to you with all my love.

When God created Adam, he whispered in his ear, ‘In all contexts of action you will recognise rules, if only the rule to grope for rules to recognise. When you cease to recognise rules, you will walk on four feet’.

Wilfrid Sellars

Introduction

Understanding, Normativity and Science

1 Understanding and Normativity

Recent work in epistemology and philosophy of science has argued that understanding is an important cognitive achievement that philosophers should seek to address for its own sake. This thesis outlines and defends a new approach that analyses the concept of understanding in relation to the phenomenon of normativity. I argue that scientific understanding, like all other forms of understanding, is constituted and sustained by the normative structure of our concrete, social world.

In contemporary philosophy, discussions of the connection between understanding and normativity have primarily arisen in relation to debates concerning the normativity of meaning, which in turn have fed into broader questions concerning the normative nature of mind, intentionality, rationality and action.[1] Here, it is argued that the meaningfulness of what we say and do is underpinned by our sensitivity to norms embedded within our activities and practices. To understand the meaning of an expression involves not simply a disposition to use that expression in a certain way, but rather a grasp of how it *ought* to be used. The crucial difference between our intelligent use of language and concepts from mere parroting or automatic behaviour is that we *understand* how what we say and do *ought* to be said and done. In that respect, our understanding reflects our ability to think, speak and act in ways that are answerable to what one *ought* to think, say, or do. This thesis is about the connection between understanding, the normativity of meaning and mind, and the capacity of us as subjects to be responsive and responsible to norms.

Before going further, however, it is necessary to clarify the concept of 'normativity'. The concept has its roots in the term 'norm', and a norm is often used in a narrow sense to refer to an explicitly formulated rule, criterion or value that serves as the basis for determining whether something is permissible or obligatory. But there is also a broader sense, according to which

[1] While not without its controversies, the thesis that the meaning of words or the contents of beliefs are normatively constituted is becoming increasingly influential. Versions have been advanced by Boghossian (1989), Cash (2008), Fennell (2013), Gibbard (1990), Ginsborg (2012), Haugeland (1998), Lance and O'Leary-Hawthorne (1997), McDowell (1994), Peregrin (2012) and Whiting (2009). A useful critical discussion of the normativity of meaning is provided by Glüer and Wilkforss (2015).

a norm is a covering term for anything that serves as a standard for success or failure of any kind. I shall understand normativity to refer to *norm-relativity* in this broader sense. From this broader perspective, normativity is a pervasive feature of human life and is found 'wherever we can speak of rules, measures, standards, exemplars, ideals, concepts […] wherever distinctions between better and worse, success and failure, can be made' (Crowell 2013: 2).

This connection between understanding and normativity in this broad sense requires further clarification. Recent work on understanding has addressed the question of whether understanding has specific *epistemic norms*; for example, whether understanding should be truth-conducive, justified or coherent.[2] But I take the connection between understanding and normativity to be more basic than this. As Joseph Rouse puts it:

> Thoughts, utterances, and actions are open to epistemic, pragmatic, moral, political and other forms of assessment, but the normativity of meaning and mind concerns their *candidacy* for assessment in these ways, not the *outcome* of those assessments. To assess the truth or falsity of a belief or statement, for example, one must understand what it says. To assess the success or failure or the moral significance of an action, one must grasp what the agent was doing or trying to do. (Rouse 2017: 546).

This suggests that there is a level of understanding and normativity that is already operative prior to the assessment of specific claims to understand something, according to which those claims are already meaningful or intelligible. I shall refer to these as *norms of intelligibility*, and this thesis shall be primarily concerned with norms of intelligibility rather than epistemic or moral norms. The central claim is that to understand means to grasp something in the light of norms of intelligibility.

In the rest of this Introduction, I provide an overview of the recent debates on the concept of understanding and highlight salient issues that I will discuss in later chapters. I will then outline my philosophical and methodological commitments before sketching out the structure of the thesis.

2 Understanding and Explanation

One of the primary reasons for the current interest in the concept of understanding is that understanding is taken to be a central goal of the sciences. The sciences strive to make sense of the world in a precise, predictive and reliable way in order to better understand it. For several authors, any attempt to account for the cognitive success of science and its achievements must therefore take the notion of understanding seriously (Elgin 2007: 34; de

[2] For example, Baumberger et al. (2016), Hills (2015), and Dellsen (2016).

Regt and Dieks 2005: 142). However, even if understanding is not as central to science as some may think, science is undoubtedly concerned with *explanation*, and there is a close conceptual connection between explanation and understanding. Explanations are one of the primary vehicles of understanding in scientific inquiry and any account of understanding should accommodate this conceptual connection. Exactly how to account for this conceptual connection, however, has been the subject of recent debate.

For much of the twentieth century, philosophers of science have been sceptical that the notion of understanding merits serious philosophical attention. Instead, they have focused their analysis on the topic of scientific explanation. Traditionally, this led to the concept of understanding being characterised in one of two ways. On the one hand, it was regarded as an idiosyncratic, psychological phenomenon that should have no bearing on the objectivity and rationality of scientific inquiry. I shall refer to this as the *psychologistic* account. On the other hand, others argued that understanding should not be treated as a psychological phenomenon, but rather as knowledge of whichever explanatory factor was favoured by a proposed model of explanation. We can call this the *explanationist* approach. Both approaches tie understanding to explanation and see it as an automatic product of an explanation. More recently, however, some philosophers have argued that the concept of understanding should be analysed for its own sake, with some degree of independence from accounts of explanation. For reasons that will become clear, accounts that take this last approach can be referred to as *manipulationist*.[3] I will now briefly review each of these approaches.

The psychologistic approach to understanding has its roots in logical positivism, and receives its paradigmatic expression in the work of Carl Hempel and his analysis of scientific explanation. Hempel (along with Paul Oppenheim) famously developed the *deductive-nomological* model of explanation (Hempel 1965). On that model, to explain something is to show how to derive the *explanandum* in a logical argument, to show that it is to be expected given our knowledge of scientific laws. Hempel's analysis of scientific explanation was carried out in formal-logical terms: scientific explanations were demarcated as an object of logical analysis and could be explicated solely in terms of an analysis of the binding relation (logical or otherwise) between the *explanans* and the *explanandum*. In doing so, the deductive-nomological model of explanation could provide an account of the objective nature of scientific inquiry (Hempel 1965: 337).

On this approach, scientific explanations could be logically analysed as objective components of scientific discourse, whereas understanding, it was argued, was of no relevance to scientific inquiry. Hempel and Oppenheim distinguished between 'understanding in the psychological sense of a feeling of empathic familiarity' from 'understanding in the

[3] The distinction between explanationism and manipulationism was proposed by Kelp (2015).

theoretical, or cognitive, sense of exhibiting the phenomenon to be explained as a special case of some general regularity' (Hempel 1965: 257). They criticised the first sense of understanding as being neither a necessary nor a sufficient requirement for scientific explanations, regarding it only in a derivative, 'pragmatic' sense that necessarily implicated the subject, thus compromising its objective status (cf. Hempel 1965: 425).[4] This is because while one may legitimately say that theory *T* explains phenomenon *P*, to speak about understanding *P* by means of *T*, one must invoke a subject. In other words, understanding involves not a two-term relation between an *explanans* (the theory) and an *explanandum* (the phenomenon), but rather a three-term relation between *explanans*, *explanandum*, and a subject (de Regt, Leonelli and Eigner 2009: 3). Although Hempel conceded that such aspects might be of interest to psychologists and sociologists studying the cognitive and psychological states involved in scientific explanation in an empirical investigation, he maintained they were irrelevant to the philosophy of science.[5] Regarding understanding in the theoretical or cognitive sense of the term, Hempel held that understanding required no more than knowledge of (approximately) correct scientific explanations, thus anticipating the explanationist approach to understanding.

The psychologistic approach towards scientific understanding is today defended by (amongst others) J.D. Trout (2002, 2007). Trout argues that understanding refers to the *sense* or *feeling* of understanding (what is referred to as the 'Aha!' moment) and as such sees it as a pernicious and highly unreliable cognitive phenomenon. For Trout, the feeling of understanding has no epistemic weight and consequently should have no bearing on the objective nature of scientific inquiry, because what makes an explanation good (or epistemically reliable) is not determined by the feeling of understanding that it produces, but rather 'concerns a property that the explanation has independent of the psychology of the explainers; it concerns features of external objects, independent of particular minds' (Trout 2002: 217). Therefore, Trout argues that philosophers of science should not resort to subject-relative features like understanding or intelligibility in order to provide an account of scientific

[4] By the 'pragmatic' aspects of explanations, Hempel and Oppenheim meant notions that refer to and describe a subject who understands, something that can potentially vary from individual to individual.

[5] This view had early critics, such as Michael Scriven (1962) and Peter Achinstein (1983). Scriven argued that a focus on the logical analysis of explanation produced an idealized model that does not bear out in practice, proposing instead a Wittgensteinian conception of explanation and understanding, which was sensitive to the context and background knowledge implicit in the scientific inquiry. Achinstein placed the pragmatic aspects of explanation centre stage, contrary to Hempel and the logical empiricist tradition, and provided a detailed analysis of the contextual relation between explanation and understanding. However, these challenges by Achinstein and Scriven did not receive the same level of attention as did later unificationist and causal-mechanist models of explanation, and consequently have not had as much impact as they deserved.

explanation. Such features, Trout claims, can be subject to both overconfidence and hindsight bias, producing what Petri Ylikoski calls the 'illusion of depth of understanding' (Ylikoski 2009), whereby scientists may be unjustifiably confident in a particular explanation or its potential for an avenue of inquiry. An account of scientific explanation that involves reference to understanding is, as a result, misguided insofar as it cannot offer a reliable epistemic cue to the value of particular explanations. What makes an explanation good is its ability to correctly identify and explain features of a given phenomenon, independently of the subject.

Explanationist approaches to understanding have a more robust characterisation of understanding and were among the first to criticise the psychologistic conception. Michael Friedman, for example, challenged the deductive-nomological model of explanation precisely because it failed to account for the features of scientific explanation that made understanding possible, rejecting the easy equation made by Hempel of understanding with the psychological aspects of explanation. Any account of scientific explanation, Friedman argued, 'should somehow connect explanation and understanding – it should tell us what kind of understanding scientific explanations provide and how they provide it' (Friedman 1974: 9). Friedman claimed that scientific understanding should not be relegated to being a merely subjective afterthought, but rather should be open to objective analysis as well. Nevertheless, Friedman shared certain assumptions with Hempel concerning what we should expect from such an account: 'what counts as an explanation should not depend on the idiosyncrasies and changing tastes of scientists and historical periods. It should not depend on such nonrational factors as which phenomena one finds somehow more natural, intelligible, or self-explanatory than others' (ibid: 14).

Friedman proceeded to argue that objective understanding is provided by explanations that *unify* more phenomena than explanations that do not. On this unificationist model, scientific understanding results from being able to 'reduce the total number of independent phenomena that we have to accept as ultimate or given' (ibid: 15).[6] In contrast to the logical methods of the deductive-nomological model, Friedman's unificationist model of explanation proposed that explanation and understanding operated at a global level insofar as current scientific knowledge was able to present a unified picture of the world. A unificationist account has

[6] To use Friedman's example, the kinetic theory of gases explains phenomena involving the behaviour of gases, such as the fact that gases approximately obey the Boyle-Charles law, by reference to the behaviour of the molecules of which gases are composed. But Friedman argues that this alone does not add anything to the understanding: it simply explains one fact in terms of another. The application of the theory does not stop here. The kinetic theory also allows us to derive other phenomena involving the behaviour of gases, such as the fact that they obey Graham's law of diffusion or that they have the specific heat capacities, from the laws of mechanics. Friedman's point here is that what the kinetic theory does is to provide an account of how independent facts are explainable via one theory (Friedman 1974: 14-15).

obvious appeal: if a set of previously distinct or disparate theories dealing with specific domains of phenomena can be unified under a single general theory, pattern or principle, then it seems obvious to say that we will have greater amount of understanding.[7] Philip Kitcher (1981, 1989), motivated in part by problems in Friedman's own account, further expanded and substantiated this unificationist model of explanation. What characterises both accounts is that their notions of scientific understanding were still modelled on the features and virtues of their particular theory of explanation. Only insofar as scientific explanation provides a unifying account of a range of different phenomena (or in Kitcher's case, *explanatory schemata* that can be applied as widely as possible) do we advance our understanding of science. Moreover, the ability of an explanation to unify disparate phenomena would also seem to capture what is going on in the 'Aha!' moment. If a particular odd-looking phenomenon is just a case of something more general, then it seems natural to suppose that we have understood something about it (Godfrey-Smith 2003: 196).

Another variant of the explanationist approach comes from the causal-mechanical model. Developed primarily by Wesley Salmon (1984), with a more recent alternative version proposed by James Woodward (2003), the causal-mechanical account argues that it is knowledge of the causes of scientific phenomena that holds the key to our understanding of the world (Salmon 1984: 260).[8] This fits with our intuitions that science does not only describe or fit phenomena into a particular theoretical framework, but that it is an explanatory virtue of scientific understanding to be able to explain why a particular phenomenon has occurred. It would seem that a causal story amounts to the right way to go about achieving this. On the causal-mechanical model of explanation, knowledge of causal-mechanisms advances our understanding because 'causal processes, causal interactions, and causal laws provide the mechanisms by which the world works; to understand why certain things happen, we need to see how they are produced by these mechanisms' (Salmon 1984: 132, quoted in de Regt, Leonelli and Eigner 2009).

The basic thrust of the explanationist approach to understanding has been recently formalised by Kareem Khalifa (2012, 2013). Rather than identifying explanatory understanding with knowledge of a particular explanatory feature (laws, causes, or unifying principles), Khalifa reconfigures the explanationist approach into a general deflationary principle:

[7] A unificationist conception of understanding has recently been defended by Gijsbers (2013).

[8] Accounts that emphasise our knowledge of mechanisms, rather than causal processes, as enabling understanding can be found in Bechtel and Richardson (1993), Machamer, Darden and Craver (2000), and Craver and Darden (2013).

> (Explanatory Model of Understanding) Any philosophically relevant ideas about scientific understanding can be captured by philosophical ideas about the epistemology of scientific explanation without loss. (Khalifa 2012: 17)

With this model, Khalifa captures the reductive connection between explanatory content and understanding without having to endorse any particular model of explanation. Explanatory understanding can simply be identified with (a) knowing that the *explanans* is true, (b) knowing that the *explanandum* is true, and (c) knowing what the correct explanatory link is between the *explanans* and *explanandum* (ibid: 26). On this view, explanatory understanding does not involve anything over and above the knowledge of the explanatory information contained within the *explanans* and the *explanans-explanandum* relation.

Manipulationists challenge both the psychologistic and explanationist approaches. Against psychologistic accounts, they hold that understanding should not be reduced to a psychological phenomenon relative to different individuals. Against explanationist approaches, they argue (i) that it is possible to acquire scientific understanding without an explanation, and (ii) that understanding cannot be reduced to knowledge of explanatory information. Manipulationist accounts do not completely divorce understanding from explanation, but they do argue that in framing understanding solely in terms of explanation, psychologistic and explanationist accounts miss important features of understanding.

First, Peter Lipton (2009) has argued that it is possible to acquire understanding of a phenomenon without requiring an explanation. To show this, Lipton makes a distinction between an explanation on the one hand and the cognitive benefits of an explanation on the other. Lipton then proceeds to identify understanding why a phenomenon occurs with the extra cognitive benefits that a good explanation provides rather than simply with the explanation itself. Such benefits, Lipton proposes, include four kinds of knowledge: (i) of causes, (ii) of necessity, (iii) of possibility, and (iv) of unification. Lipton then outlines ways in which the above four cognitive benefits might also be acquired by routes that do not pass through explanations. To illustrate Lipton's argument, we can focus on one of his examples: causation.

Lipton argues that causal knowledge can be provided by means other than explanation. Experimentation, observation, manipulation and inference are all plausible candidates for being sources of causal knowledge that are distinct from providing an explanation. However, this does not yet illustrate how understanding can be gained without going through an explanation: 'We need cases that, in addition to not being explanations themselves, do not work by means of generating explanations that are then the proximate cause of the consequent understanding' (ibid: 45). The worry is that experiment and other sources are only one stage in developing an explanation and this by itself does not facilitate understanding.

Yet, Lipton suggests, this suspicion only seems compelling if we focus on causal information that takes the form of propositions explicitly entertained. The suspicion loses its

force, however, if we look to cases where what is produced is not explicitly propositional explanations but a tacit, pre-articulate familiarity with things. As an example, Lipton cites the kind of tacit knowledge we might gain through the use of images and models. Coming to understand 'the why' of retrograde motion, for instance, may not be possible until it has been demonstrated visually in a planetarium. In such cases, 'visual devices may convey causal information without recourse to explanation' (ibid: 45). Furthermore, such tacit capabilities should also make us question the idea that understanding must be explicable in general propositions or rules:

> […] a scientist may gain a sophisticated understanding of the behaviour of a complicated piece of machinery by becoming an expert at using it, and that understanding consists in part in the acquisition of causal information that the scientist may be in no position to articulate. (ibid: 45)

If Lipton is correct, then explanation is not a necessary condition for scientific understanding. However, even if this is the case, Lipton still leans towards identifying understanding with a specific kind of knowledge (cf. Lipton 1991) – for example, knowledge of causes – whereas other manipulationist accounts have disputed this, as we will see in the following section.

The second general argument for the manipulationist challenge is that understanding is not reducible to knowledge of explanations. Instead, it is argued that subjects must have additional cognitive abilities to count as understanding. The abilities allow subjects to *manipulate* the explanatory information in some way. For Henk de Regt (2009a), it is the possession of specific *epistemic skills* that enables an individual to extract the right information from an explanation in order to acquire understanding: 'The extra ingredient needed to construct the explanation is a *skill*: the ability to construct deductive arguments from the available knowledge' (de Regt 2009a: 26). Alternatively, Stephen Grimm (2010: 340–1) argues that explanatory understanding requires that a subject S can anticipate how changes in one variable in a causal sequence can lead to changes in another variable; the key here is the ability to answer 'what-if-things-had-been-different' questions. He also requires S to be able to apply general expressions of these causal relations to particular cases. Thus, Grimm requires S to be able to make predictions for new variations on previously studied examples.[9] For these accounts, therefore, it is possible to have an explanation without understanding the phenomenon being explained, because understanding depends upon the possession and development of certain cognitive abilities. Although one may have a correct explanation of a

[9] Grimm adapts this from Woodward's (2003) manipulationist account of causal explanation, according to which the measure of understanding is the ability to answer 'What-if-things-were-different?' questions. A similar view is defended in the epistemological literature by Alison Hills (2015).

phenomenon, there will be no understanding if the subject does not have the requisite cognitive abilities. As such, manipulationist accounts propose that understanding should be considered independently of existing accounts of scientific explanation.

The dispute between the psychologistic, explanationist and manipulationist accounts of understanding indicate that there is no agreed upon way of characterising explanatory understanding. In my view, the psychologistic approach is deeply inadequate. I will argue that it is a mistake to think of understanding as an internal, mental phenomenon. Like the explanationists and manipulationists, I hold that understanding cannot be reduced to an idiosyncratic 'feeling'. First, psychologistic accounts fail to capture the fact that our understanding can and should be accountable and responsive to the world. It would be a mistake to identify understanding solely with the phenomenological feeling that may accompany it. Second, psychologistic approaches reflect a wholly inaccurate conception of what it means to be a subject of understanding. Unlike explanationist accounts, I do not think that understanding can or should be reduced to mere knowledge of explanations. Not only does this fail to do justice to the range of ways in which we actually make sense of the world, it also relies on a limited conception of the role of explanatory discourse in scientific practice. As I shall argue, explanations do more than simply provide us with more knowledge of the world. Within scientific communities and disciplines, explanations fulfil an important function of coordinating scientific inquiry in relations to the aims and goals of particular disciplines. In that respect, I will argue that making sense of explanatory understanding in scientific practice requires attention to the *social* function of explanations. Finally, although my own account comes closest to manipulationist approaches, I will argue that the abilities and skills proposed by manipulationists need to be developed in line with a broader conception of how understanding is situated within human and scientific practices.

The upshot of these debates is that the ties between understanding and explanation have been gradually disentangled without being severed completely. Nevertheless, explanations are an essential vehicle of scientific understanding. If understanding does not automatically follow from having an explanation, as the manipulationists hold, then these arguments raise the question of how understanding is connected to explanation and what the nature of that connection is. I address this question in Chapter Five. Furthermore, if, as the explanationists hold, understanding amounts to knowledge of explanations, then it is also important to address the connection between understanding and knowledge.

3 Understanding and Knowledge

In philosophy of science, the debates concerning understanding have primarily, but not exclusively, focused on the relation between understanding and explanation. Within epistemology, the debate has primarily focused on the relations between understanding and knowledge. Although it is a matter of some debate what exactly the epistemic conditions on understanding are, it is widely assumed that those conditions can be fleshed out in similar terms to the traditional conditions of knowledge: belief, truth, and justification.[10] As the focus of this thesis is scientific understanding, I shall not touch on as many details of the epistemological debates on understanding as I did with explanation. However, we see in the epistemological literature a similar move where understanding is not easily identified with knowledge of the world. Instead, understanding seems to introduce additional conditions and features. As such, understanding and knowledge can also be seen to have some degree of independence from one another.

First, it is a commonplace that understanding requires more than believing or accepting or even knowing isolated pieces of information. Epistemologists have argued that understanding requires that we also 'see' or 'grasp' how those bits of information fit together. Thus, Wayne Riggs argues that '[a]n important difference between merely believing a bunch of true statements within subject matter M, and having understanding of M, is that one somehow sees the way things fit together. There is a pattern discerned within all the individual bits of information or knowledge' (Riggs 2003: 218). In a similar vein, Catherine Elgin suggests that 'understanding a particular fact or finding, concept or value, technique or law is largely a matter of knowing where it fits and how it functions in the matrix of commitments that constitute the science' (Elgin 1993: 14-15). Finally, Jonathan Kvanvig argues that understanding involves, 'grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information' (Kvanvig 2003: 192).

The claim that understanding involves grasping or seeing how various bits of information relate to one another has resulted in analyses of a particular type of understanding: *objectual understanding*. Objectual understanding is understanding a thing, a domain of things, or a subject matter. An example of objectual understanding is discussed by Christoph Baumberger:

> Understanding global warming involves, for instance, understanding what effects (on natural and social systems) it will have, how it is linked to human activities (such as burning fossil

[10] Recent contributions to the epistemological literature on understanding include: Baumberger (2011), Elgin (2004, 2007), Grimm (2006, 2012), Grimm, Baumberger and Ammon (2016), Hills (2015), Kelp (2015), Kvanvig (2003), Riggs (2003), and Zagzebski (2001). A useful overview of the current issues is Baumberger et al. (2016).

> fuels and deforestation) and related phenomena (such as the destruction of stratosphere ozone and global dimming), how far greenhouse gas emissions and, as a result, temperatures are likely to rise in the future, and how the changes will vary over the globe. (Baumberger 2011: 77-78)

The general idea then is that while understanding is related to knowledge, it is typically presented as a higher-level form of cognition insofar as it is taken to require the synthesis and organisation of our beliefs, rather than simply the possession of them. Understanding in this respect involves the ability to appreciate how different things in a body of knowledge relate to one another and 'fit together'.

A second reason why understanding and knowledge seem to be different is that understanding comes in degrees and can vary in breadth and depth. We often describe some as having a better understanding of a subject matter than others. Rather than withholding attributions of understanding from, say, a student, altogether, it might seem more plausible to say that they have *some* understanding, and the teacher has *better* understanding. This seems to be in contrast to standard analyses of knowledge, which suggest that knowledge is an all-or-nothing affair: you either know something or you do not. Instead, understanding can be partial and come in degrees.

This leads to the third aspect, which concerns whether factivity – the truth of our beliefs – is a condition of understanding. Some maintain that truth and factivity are genuine conditions on understanding in exactly the same way as they are on knowledge (Kvanvig 2003; Kelp 2015). Others have argued that a strong factivity condition seems too demanding, particularly for objectual understanding. Here, it is argued that as long as there is a core of true beliefs concerning a subject matter, a few falsehoods at the periphery may be tolerated, without denying someone understanding (Kvanvig 2009). One of the key arguments against a strong factivity condition comes from the history and practice of science. The claim here is that a factive condition on scientific understanding is too restrictive: it would mean denying understanding to many central components of scientific thinking (Elgin 2007; De Regt 2015). Elgin, for example, draws on the following case:

> A central tenet of Copernicus's theory is the contention that the Earth travels around the sun in a circular orbit. Kepler improved on Copernicus by contending that the Earth's orbit is not circular, but elliptical […] Despite the fact that Copernicus's central claim was strictly false, the theory it belongs to constitutes a major advance in understanding over the Ptolemaic theory it replaced. Kepler's theory is a further advance in understanding, and the current theory is yet a further advance…With each step in the sequence, we understand the motion of the planets better than we did before. But no one claims that science has as yet arrived at the truth about the motion of the planets. Should we say that the use of the term 'understanding' that applies to such cases should be of no interest to epistemology? (Elgin 2007: 37-8)

The issue that Elgin is raising is that growth of understanding seems to be a natural feature of scientific progress: we can say that successive theories, even if now known to be wrong, constituted advances in our understanding. In that respect, it would seem possible to attribute some degree of understanding to past theories and scientists. Whether you agree with this or not depends on how strong a factivity condition understanding ought to have.

A similar point concerns the use of idealized models as a key tool in scientific investigation. Strictly speaking, idealized, theoretical models are widely known not to be true about the world. Indeed, as literal descriptions of the world, they are false. The model underlying the ideal gas law, for example, accounts for the behaviour of real gases by assuming the latter consist of perfectly elastic point masses that do not interact with each other. Nevertheless, the idealizing assumptions out of which the model is constructed seem to facilitate scientific understanding: the model has enabled substantial scientific research and many chemistry textbooks begin with discussions of the ideal gas model and law, even though it would appear not to satisfy a strong factivity condition.

Furthermore, scientists regularly use different models for the same phenomenon despite them being inconsistent with one another. For instance, there are a range of models of the atom – the liquid drop model, the MIT bag model, and the shell model – that incorporate known falsehoods concerning their target (Morrison 2011). Each model provides some form of understanding of the atom, even though taken together they would present a very inconsistent picture. Finally, many models used within scientific practice are non-propositional representations, such as diagrams, maps, graphs, and three-dimensional models. If the content of such representations is not fully explicable in terms of propositions, the resulting understanding cannot be even moderately factive (Riggs 2003: 218).

In this thesis, I hold that understanding is not just more knowledge about the world, or that it can be entirely explicated in terms of true or accurate propositions about the world. However, my reasons for doing so depart from current accounts in epistemology because I argue that understanding is itself a prerequisite for having knowledge in the first place.[11] I discuss this point in more detail in Chapters One and Two. Like the above recent developments in the epistemology of understanding suggest, my own account incorporates the idea that understanding has a holistic character and comes in varying degrees. With regards to the factivity of understanding, my own contention is that a strong factivity condition is indeed too demanding for an adequate conception of understanding in general and in accounts of scientific understanding specifically. Denying strong factivity, however, does not mean that

[11] This idea was neatly expressed by John Haugeland in an unfinished manuscript: 'One cannot have a cognitive attitude toward a proposition that one does not understand, but one can certainly understand a proposition without having any attitude toward it at all' (Haugeland 2017: 303).

our understanding should not be accountable to how the world is. What I am interested in is what it means for understanding to be accountable to anything in the first place.

4 Methodological Commitments

The account of understanding proposed in this work has its roots in three broad schools of thought. Each one thinks that only by beginning with discursive human *practices* can we make sense of how our understanding can be directed towards and responsive to the world. I share this methodological commitment and defend the claim that human understanding cannot be adequately understood as a narrowly psychological and subjective phenomenon. Instead, the central idea I take from each school of thought is that understanding and normativity are phenomena that are ineliminably grounded in social, embodied practices.

The first school of thought derives from the phenomenological tradition, inaugurated by Edmund Husserl and substantially developed by Martin Heidegger and Maurice Merleau-Ponty. Phenomenology is a diffuse branch of philosophy, and there have always been important disagreements between thinkers in this tradition over the exact nature, meaning and method of phenomenology.[12] Although phenomenological approaches typically encompass a wide range of characteristics of human life, including perception, intersubjectivity and temporality, a major, common concern of phenomenological investigations is the constitution of *meaning*. That is, broadly speaking, phenomenology is concerned with how we make sense of the world, and these various characteristics are tied together by the fact that each plays some role in constituting the way in which our world is meaningful.

Although earlier forms of phenomenology were primarily concerned with uncovering necessary and invariant structures of meaning, contemporary phenomenological work typically eschews the search for transcendent principles in favour of focusing on embodied practice as the *de facto* site of human meaning. Phenomenological thinkers have been brought into productive dialogue with Anglo-American philosophy of mind, epistemology and language, for example in the work of Hubert Dreyfus (1991), Charles Taylor (1985) and John Haugeland (1998). Here we see an emphasis on skilful and practical agency, local and contextual epistemic practices, and the role of social relations as key features of our understanding of the world. More recent developments have seen the application of phenomenological ideas to cognitive science, particularly in the conception of human cognition as embodied, extended and embedded.[13] In Chapter One, I draw on Heidegger's

[12] For a useful discussion of these disagreements, see Crowell (2002) and Fernandez (2017).
[13] For example, in Noë (2012), Clark (2011), and Thompson (2010).

account of what it means to be a subject of understanding which will form the foundation for the rest of the thesis.

Second, I draw on what has come to be known as the 'Pittsburgh School' of contemporary neo-pragmatism (Maher 2012), represented paradigmatically by Wilfrid Sellars, John McDowell, Robert Brandom, and John Haugeland. Each of these thinkers can be characterised by a commitment to anti-reductionism and anti-representationalism in the philosophy of mind and epistemology, and have their philosophical roots in Wittgenstein. These philosophers are committed to the principle that the best place from which to begin thinking about conceptual understanding is by foregrounding our practical interactions with the world and with others, paradigmatically in our conceptually mediated linguistic practices. Each takes conceptual normativity to be a fundamental aspect of how we make sense of the world, developing Sellars' provocative claim that as rational creatures – that is, beings capable of language, thought, intention, and normative accountability – we are situated within a 'space of reasons' set over against a 'space of mere causes': 'in characterising an episode or state; we are placing it in *the logical space of reasons* of *justifying and being able to justify* what one says' (Sellars 1997: 76). For Sellars, it would be a mistake to characterise states such as beliefs, their corresponding propositional content, and knowledge more generally, in terms of mere empirical descriptions. Accordingly, the space of reasons is not the space in which we describe the psychology of the acts of thinking, but the space in which the normative conditions of our thoughts and actions are at stake. Sellars famously argued that intentional mental states are best understood as derivative and dependent upon meaningful, discursive practice. This commitment to the explanatory priority of the pragmatic is shared by his intellectual descendants, in which the order of explanation shifts from discursive, normative use, through to conceptual content, and to mind. In Chapter Two, I build upon Brandom's 'normative inferentialism' to explicate the idea that understanding is constituted by our inferential ability governed by material norms of inference.

The third school of thought that I draw upon is the philosophy of scientific practice. In many of its traditional forms, philosophy of science tended to reflect primarily on the relation between scientific theories and the world, oftentimes to the neglect of scientific practice. This theoretical orientation was substantially challenged in the latter half of the twentieth century, with on the one hand, the emergence of detailed, empirically-informed case studies of sociological aspects of scientific research, and on the other hand, a shift towards scientific experiment and a focus on the role of instrumentation and technology in the development of

scientific knowledge and understanding.[14] A significant motivation for focusing on the philosophy of scientific practice is that it prompts reconsideration of how to approach issues in the philosophy of science. Traditional epistemological questions such as the nature of truth, fact, observation, explanation, justification and evidence become 'usefully re-framed in terms of activities' (Ankeny et al. 2011: 305).

The shift towards scientific practice was heralded by the 'New Experimentalism' (Ackermann 1989): the view that experimental work encompasses much more than just the testing of high-level theories. Instead, scientific experiment incorporates a range of epistemic activities that function relatively independently of an explanatory, higher-order theory, which in turn involve a wide assemblage of skilful performance, practical ability, and material intervention.

The initial emphasis on the practical over the theoretical has tempered in recent years, and philosophers of scientific practice have provided careful analysis of theoretical activities, such as modelling, mathematics, and explanatory reasoning. The focus on scientific practice should therefore not be read as introducing a pernicious dichotomy between theory and practice, but instead as a commitment to provide a more descriptively accurate and functional account of how the diverse aspects of the scientific enterprise are utilised in ongoing research. The focus is less on the clean and tidy representation of scientific work as presented in scientific publications, but on the messy, concrete business of doing science. The methodological focus on scientific practice encapsulates a broader concern for the *processes* of scientific inquiry, rather than just the *products*.

Andrea Woody (2014: 123-124) provides a useful schematic outline of the broad shifts away from more traditional philosophy of science that characterise the turn to scientific practice. First, we see a move away from treating theories as formalisable conceptual structures. Instead, scientific theories are understood to be embedded in a range of representational artefacts – graphs, diagrams, equations, models, photographs, instrumental inscriptions, written reports, and computer simulations – which form the basis for theoretical work, and which, like experimental practice, is 'shaped by the practical concerns and contingent, contextually determined goals' (Woody 2014: 124). Second, there has been a shift away from *a priori* analysis of the rationality and justification of scientific beliefs and concepts, and towards a more descriptively adequate examination of the reasoning invoked by scientists in particular contexts (ibid: 123). The third shift lies in no longer treating scientists as ideal, disembodied and rational agents, but as embodied human practitioners

[14] Earlier works which exemplified this shift to scientific practice include Latour and Woolgar (1979), Hacking (1983), Pickering (1995), Franklin (1989), Galison (1987, 1997), Rouse (1987), Traweek (1992), Kohler (1994), and Rheinberger (1997).

situated in a world shaped by their training, expertise, and characterised by variations in epistemic, social, and political values (e.g., Longino 1992, 2001). One prominent result of this shift has been a turn to the resources of cognitive science and psychology as a way of modelling scientists' online and offline cognitive processes as a way of informing our understanding of scientific practice itself (e.g., Nersessian 2008). Fourth, Woody argues that we also see a shift away from an individualistic orientation in epistemology and towards social epistemological questions. This encompasses questions concerning the role of scientific communities in constituting and sustaining knowledge, and questions concerning the various ways in which knowledge is shaped by broader institutional, political and societal concerns and mitigated by questions of authority, power and trust.

In Chapters Four and Five, I situate my account of understanding in general, and scientific understanding more specifically, in relation to the philosophy of scientific practice in two ways: through an examination of model-based understanding and explanatory understanding respectively. Each chapter draws on recent practice-based accounts of models and scientific explanations. The focus here is an analysis of how models and explanations function as vehicles and tools for understanding, and how they are situated in the practice of ongoing research.

Combining these different philosophical traditions and schools of thought to address questions in contemporary epistemology and philosophy of science undoubtedly presents a challenge. The issues, conceptual resources and language of each approach are not easily integrated with one another. However, while the problems and issues addressed by these different schools of thought may be couched in different and possibly unfamiliar terms, I believe that there are substantial commonalities and shared concerns. In that respect, this thesis owes much to the work of Joseph Rouse (1987, 2002, 2015), who has continually argued for the benefits of incorporating these different traditions into one framework and whose arguments have inspired much of the content of this thesis. My aim in this thesis, however, is to show how this framework can be applied to contemporary philosophical analysis of the concept of understanding in epistemology and philosophy of science, something that Rouse does not explicitly address. To that end, I shall conclude this Introduction with a brief overview of the thesis.

5 Overview

The thesis is divided into two parts, reflecting the division of labour that has emerged within philosophical accounts of understanding. Epistemological approaches have typically focused on the subject of understanding: what characteristics a subject must have and what they must

be capable of doing in order to count as understanding. Within the philosophy of science, by contrast, the focus has been on how the vehicles of understanding, such as theories, explanations and models, facilitate scientific understanding. Thus, Part I (Chapters One to Three) of the thesis addresses the question of the nature of the subject of understanding and Part II (Chapters Four and Five) applies the account developed there to the vehicles of scientific understanding and their function within scientific practice.

As we have seen, the nature of what it means to be a subject of understanding has so far been framed in terms of standard epistemological concerns. On the one hand, characterising the subject of understanding has revolved around the question of the extent to which understanding involves the traditional constituents of knowledge. On the other hand, the subjective nature of understanding has been cast in inherently negative terms as a narrowly psychological phenomenon that has no bearing on the objectivity or rationality of scientific research, as illustrated by the psychologistic approaches in Section Two. Both approaches, I suggest, are inadequate. Chapter One proposes an alternative starting point, one that begins within the phenomenological tradition and specifically with the work of Heidegger. In this chapter, I argue that Heidegger presents a richer, phenomenological conception of what it means to be a subject of understanding. For Heidegger, to be a subject of understanding is to be situated in a world of meaning and intelligibility, shaped by one's place in social, norm-governed human practices. From this phenomenological perspective, understanding is a practical, normative capacity to deal appropriately with one's surroundings. I argue that the normative nature of understanding is essential to Heidegger's characterisation of the subject of understanding: a necessary condition on being a subject of understanding is to be capable of responding to norms *as* norms.

Chapter Two develops this Heideggerian starting point by turning to Robert Brandom's normative inferentialism. Several authors have argued that inferential ability is a constitutive criterion of understanding. We find such a commitment, for example, in manipulationist accounts and in the notion of 'grasping'. While broadly agreeing with this claim, I use Brandom's work to show that this inferential ability cannot be detached from our norm-governed practices: for Brandom, semantic and intentional content is conferred on our conceptual understanding by being situated in the 'game of giving and asking for reasons' (Brandom 1994). The focus of Brandom's account is conceptual understanding: our understanding of semantic meaning and conceptual content. It reflects a concern with the fact that understanding does not merely allow us to *do* things in the world, but to *say* something about it too. Brandom's inferentialism shows that understanding *qua* inferential ability should be analysed in terms of its role in our discursive, normative practices.

In Chapter Three, I focus more directly on the nature of normativity and its relation to human practices. I review two prominent accounts of normativity and argue that both are

inadequate. First, I show how a *norms as rules* approach is subject to the Wittgensteinian regress of rules problem. Second, I argue against a *norms as regularities* approach, on the grounds that it fails to instantiate genuine normativity. As an alternative to these accounts, I propose that norms are sustained by the *normative attitudes* of participants within a practice through relations of mutual accountability between them. Normativity is thus understood to be inherently social. Although the concept of understanding is of less direct focus in this chapter, I argue that focusing on the nature of normativity and its manifestation in human practices points to the essentially social nature of understanding.

Chapter Four shifts away from the subject of understanding and turns to the vehicles of understanding in scientific practice. The focus of this chapter is model-based understanding. The construction and analysis of models is a ubiquitous feature of contemporary scientific research. Current accounts of model-based understanding and knowledge have so far focused on the issue of the use of idealizations in building models and how this affects the 'factivity' of scientific understanding. I develop an alternative, inferentialist approach to models to capture how models facilitate understanding in scientific practice. Prior to the assessment of the representational accuracy of models, I argue that models are tools for 'conceptual articulation', in the sense that they serve to facilitate scientific understanding by articulating the meaning and inferential consequences of the scientific concepts out of which they are built. As such, models are tools for rendering the world and the phenomena they investigate intelligible in the first place. I conclude that the issues of representational accuracy and factivity can only be answered once we know what models enable us to say about the world.

In Chapter Five, I focus on the role of explanations in facilitating scientific understanding. A central debate within contemporary debates on understanding is whether explanatory understanding is adequately characterised as knowledge of explanations, or whether it involves further cognitive abilities in relation to that knowledge. I address two issues in this chapter. First, I identify and challenge an assumption that underpins the contemporary debate, which presumes that the proper level of focus for explanatory scientific understanding is at the level of the *individual subject*. I argue that this assumption fails to capture the diversity and role of explanations in scientific practice. Building on Andrea Woody's recent 'functional' approach to scientific explanation and explanatory discourse, I argue that scientific explanatory understanding is inherently *social*, and that the function of explanations is to sculpt and make explicit the norms of intelligibility required for a scientific understanding of the world. Second, I argue against recent objections to the role of skills in scientific understanding by appealing to John Haugeland's account of the constitutive normativity of scientific understanding. I argue that Woody's and Haugeland's account complement one another by reaching a similar conclusion: that explanatory understanding involves grasping phenomena in the light of constitutive norms and standards of a scientific domain.

The central claim of this thesis is that understanding in general and scientific understanding in particular are inherently norm-governed phenomena that are unintelligible without reference to the normativity embedded in our social and scientific practices. By framing understanding in terms of normativity, I show that current analyses of understanding are too restrictive and fail to capture the connection between our understanding and human, social practices.

Chapter One

The Subject of Understanding

1 Introduction

Recent work in epistemology and philosophy of science has argued that understanding is an important cognitive state that philosophers should seek to analyse. This chapter introduces a new perspective on understanding by drawing on the phenomenological tradition, particularly the work of Martin Heidegger. Central to Heidegger's characterisation of human understanding is its ontological character: to understand an entity is always to understand the ways in which that entity can *be*. Furthermore, I argue that Heidegger's conception of understanding is distinctive in virtue of the fact that, unlike contemporary approaches to understanding, he is primarily interested in what it means to be a *subject of understanding*. The aim of this chapter is to show how, from a Heideggerian perspective, our understanding of what things *are* cannot be made sense of without reference to what it means to *be* a subject of understanding.

The discussion will proceed as follows. Section Two introduces a central point of contention in current debates on understanding, which is whether understanding should be characterised as a particular kind of cognitive state or as an ability. Although I will argue that understanding is more appropriately characterised as an ability, the purpose of this section is to set up Heidegger's account and to situate it in relation to other perspectives on understanding. Section Three discusses Heidegger's conception of ontological understanding. In Section Four, I show how and why Heidegger's conception of understanding has its place in his broader investigation of what it means for 'Dasein' - Heidegger's term for the subject who understands - *to be*. Section Five concludes by considering some implications of a Heideggerian conception of human understanding for an account of scientific understanding.

2 Understanding as an Ability

A central question in recent debates on understanding is whether understanding is appropriately characterised as a kind of psychological state or as a kind of ability. This section will present some arguments for the ability-view of understanding. The primary purpose of this section is to motivate Heidegger's own conception of understanding and to locate Heidegger's approach in relation to other perspectives on understanding.

In contrast to explanation, understanding has often been felt to be too subjective to merit serious discussion by philosophers of science. One reason for this is a tendency to identify understanding with the distinctive *Aha!* feeling that we often experience when something is explained or makes sense. On this view, understanding is no more than the *feeling* or *sense* of understanding, a psychological by-product that is idiosyncratic to different individuals. Thus, Carl Hempel argued:

> Very broadly speaking, to explain something to a person is to make it plain and intelligible to him, to make him understand it. Thus construed, the word "explanation" and its cognates are pragmatic terms: their use requires reference to the persons involved in the process of explaining. In a pragmatic context we might say, for example, that a given account *A* explains fact *X* to person P_1. We will then have to bear in mind that the same account may well not constitute an explanation of *X* for another person P_2, who might not even regard *X* as requiring an explanation, or who might find the account *A* unintelligible or unilluminating, or irrelevant to what puzzles him about *X*. Explanation in this pragmatic sense is thus a relative notion: something can be significantly said to constitute an explanation in this sense only for this or that individual. (Hempel 1965: 425-26)

This Hempelian view of understanding has also been recently defended by J.D. Trout (2002), who argues that: 'What makes an explanation good concerns a property that it has independent of the psychology of the explainers; it concerns features of external objects, independent of particular minds' (Trout 2002: 217). For both Hempel and Trout, understanding is assumed to be epistemically irrelevant to the philosophy of science. Not only does it seem to introduce a form of relativism, the feeling or sense of understanding could be epistemically vicious, leading to overconfidence or bias in favour of one's hypotheses or explanations.[1]

However, the claim that understanding is no more than a psychological state or feeling has been criticised by several authors. Some, like Henk de Regt (2009a), have argued that while Hempel was right to treat understanding as a pragmatic concept, it does not follow from this that the notion of understanding is philosophically or epistemically irrelevant. Others, like

[1] However, Lipton (2009) contends that the sense of understanding can be epistemically beneficial to scientific inquiry, claiming that the *Aha!* experience might plausibly work not only as an incentive, but also as grounds for scientists' choice of suitable explanations.

Petri Ylikoski (2009), have argued that although the quality or adequacy of an explanation might produce a phenomenologically distinctive psychological state, this should not be conflated with the notion of scientific understanding itself. Ideally, understanding and the feeling of understanding would track each other, but in practice they rarely do. The feeling of understanding is a fallible and unreliable measure of understanding: we can sometimes have a false sense of understanding and sometimes we can understand something without any associated feeling or psychological experience (Ylikoski 2009: 103).[2] Therefore, while Hempel may have been right to be suspicious of the feeling of understanding that explanations might produce, it is not the case that this feeling should be automatically identified with the concept of understanding as such.

Another early critique of the idea that understanding is a psychological or mental state stems from Wittgenstein in the *Philosophical Investigations* (hereafter *PI*).[3] Central to these sections of the *Investigations* is an 'attempt to loosen the grip of the picture of understanding as an internal, mental state' (McGinn 1997: 91). To do this, Wittgenstein draws our attention to the differences between, on the one hand, the grammar of the concept of understanding and the concept of a conscious mental state, and on the other, the concept of understanding and the concept of an internal mechanism. Wittgenstein's remarks on understanding are a useful way of anticipating Heidegger's position, so I will spend some time considering them.

Wittgenstein suggests that it is distinctive of the grammar of concepts that describe conscious mental states – for example, the concept of being in pain, of being depressed, of hearing a distinct sound – that temporal concepts like duration, interruption and continuity, as well as concepts of intensity or degree, can all be intelligibly applied to them. We can, for instance, think of a pain being felt at a precise moment, enduring for a finite period, and gradually dissipating, or of the state that we are in when listening to a piece of music (*PI*: §154). What is important to recognise, Wittgenstein argues, is that when we attend to the grammar of the concept of understanding, the application of such temporal or qualitative concepts does not map quite so easily on to it. Thus, he writes (*PI*: 65):

> "Understanding a word": a state. But a *mental* state? – We call dejection, excitement, pain, mental states. Carry out a grammatical investigation as follows: we say
>
> "He felt dejected the whole day"
>
> "He was in great excitement the whole day"

[2] Michael Scriven, an early critic of Hempel's account of understanding, draws the same conclusion, arguing that it is absurd to identify the sense of understanding with understanding itself, since the former can be so easily mistaken (Scriven 1962: 225).

[3] Wittgenstein's primary discussion of understanding can be found in §§148-155 of the *Investigations*. Overall, Wittgenstein's analysis of the 'grammar' of understanding has not filtered through into contemporary discussions of understanding. However, there are a few exceptions. See, for instance, Kuorikoski (2011).

Could understanding be talked about in the same way? Of course, it's not impossible to substitute understanding in the above sentences, but the point is that it would sound odd to talk of having understood 'uninterruptedly since yesterday'. Analysis of linguistic usage can only take us so far, but Wittgenstein's broader point here is that if we assume from the outset that the concept of understanding denotes a conscious, mental state, this will obscure the evident differences between our various uses of the concept. As such, let us call *Cartesianism* in the theory of understanding the claim that understanding is a conscious mental state or process.[4]

Wittgenstein's alternative to Cartesianism is the suggestion that the grammar of understanding, by contrast, is not linked with the idea of anything's 'occurring in our minds' but rather connects more readily with the concept of an *ability*: we attribute understanding because of someone's ability to reliably do something and do it correctly and well (where we have 'mastered a technique' *PI*: §150). Indeed, at *PI* §150 and §182, Wittgenstein notes that the grammar of 'understands' is closely related to that of 'can' or 'is able to' or 'know how to'. McGinn suggests that this does not preclude us talking about the intensity or degree of understanding, but it should be recognised that the way in which these concepts apply is in a different sense, more akin to the level of our abilities; that is, how *well* we are able to *do* something, the extent or degree of our abilities (McGinn 1997: 91).

According to a different approach that Wittgenstein considers, the concept of understanding describes an underlying internal mechanism, which explains our abilities. We can see related ideas implicit in contemporary accounts of understanding. For example, several authors agree that to understand a phenomenon, one must be able to not only believe or recall particular facts, explanations, or theoretical principles, but must also be able to 'grasp' or 'see' the connection between them.[5] From this perspective, understanding is treated as a cognitive state, and 'grasping' and 'seeing' are taken to be 'psychological acts' (Grimm 2010: 342). It is coming to grasp dependency relations between propositions and variables within an explanation that constitutes understanding as a specific kind of cognitive achievement. Extending this line of thought, it has been argued (Grimm 2006: 531-532; Newman 2012: 15-16) that grasping involves not just beliefs about these connections between various items of knowledge, but also the ability to anticipate how a change in one of the variables leads to changes in the others.

[4] As suggested in Schear (2013: 164).

[5] For example, Elgin (2007), Grimm (2006) and Riggs (2003).

However, Wittgenstein argues that such a picture does not square with the grammar of our uses of the concept of understanding in our social, linguistic practices. The issue here is in differences of evaluable criteria: the standards by which we judge claims to understanding do not pertain to the appropriate functioning of an internal mechanism but rather to our manifest abilities and public performances, and it is this aspect that is obscured or overlooked in the focus on 'grasping'. These standards and criteria do not primarily concern private cognitive processes, brain states, or, even less, subjective feelings. Thus, no criterion for a speaker's understanding a word, or knowing how to add, needs to invoke 'the structure of the [cognitive/mental/brain] apparatus, as distinct from its effects' (*PI*: §149). Rather, as McGinn notes, 'our use of the concept of understanding is linked in complex ways with a speaker's participation in a characteristic form of life' (McGinn 1997: 92). What is judged, assessed and held accountable are our manifest performances – what we say or do in our social practices – and it is the public normativity constitutive of the grammar of understanding that Wittgenstein's remarks identify as essential to the concept of understanding as such.

Thus, for Wittgenstein, the concept of understanding 'connects, not with anything that happens at the time of being attributed, but with an ongoing pattern of performance' (ibid: 93). Such patterns of performance are to be understood as subject to criteria established within our practices, communities and forms of life. As Jaakko Kuorikoski argues, although an internal mechanism taking place in the privacy of individual minds may well function as a *causal* prerequisite for the possible fulfilment of these criteria, the mechanisms themselves are not the criteria of understanding, in the sense that we would have to know whether these processes were functioning properly in order say whether somebody *really* understands something – 'just as we do not need to know the cognitive and neural processes that enable one to ride a bike in order to judge whether one can really ride a bike' (Kuorikoski 2011: 171). The deeper, underlying point is that understanding is a *categorically* different kind of thing: one that applies to and is constituted by public standards of correctness, propriety and reliable ability. The *correctness* of the internal mechanisms is judged by the external displays of understanding, not the other way around (Ylikoski 2009).

These Wittgensteinian arguments thus suggest that understanding is not appropriately characterised as a psychological state or act, but rather as 'akin to an ability'. Before turning to Heidegger's own conception of understanding, it is worth making two final points. First, to the extent that the notion of understanding has been linked to the subject of understanding, we have seen that this has traditionally been characterised in negative terms. It is precisely because of the inherently subjective and psychologically idiosyncratic nature of understanding that it has been held to be epistemically irrelevant, particularly with regards to scientific inquiry. Second, concerning Wittgenstein's remarks on understanding, notice that his target is a particular form of philosophical *psychology*, namely Cartesianism, which treats

understanding as a conscious, mental state. Moreover, I have suggested that this Cartesianism informs some accounts within the current debates on understanding, particularly those that equate understanding with the psychological feeling of understanding or those that identify understanding with some internal mechanism, such as the psychological act of 'grasping'. However, Wittgenstein's remarks suggest that beyond misconstruing the grammar of understanding, the real problem with Cartesianism in theories of understanding is that it pictures understanding in complete abstraction from its surroundings and its place within a 'form of life'. The Heideggerian conception of understanding, I shall argue, can be seen as an attempt to address both of these issues: the negative characterisation of the understanding subject and the relation of understanding to a 'form of life'. With that in mind, I now turn to Heidegger.

3 Heidegger on Understanding

Although Heidegger is certainly keen to dispel the image of understanding as a conscious, mental state in *Being and Time* (Heidegger 2008 [1962]; hereafter *BT*), it is not his primary target. As Joseph Schear notes, Heidegger's resistance to Cartesianism in philosophical psychology is 'eclipsed by, or rather finds its place within, Heidegger's resistance to Cartesianism *in ontology*' (Schear 2013: 170). What this means is that Heidegger is less concerned with attacking the Cartesian idea that understanding is a conscious mental state, and more concerned with the Cartesian conception of the *subject* of understanding that the psychological view presupposes. According to this conception, the subject is a *res*, a special kind of *thing* – an entity whose mode of being is what Heidegger calls *Vorhandensein*. As we shall see, Heidegger rejects this Cartesian conception of the subject and offers a richer, more phenomenologically-informed alternative in its place. The overall aim of this section is to introduce Heidegger's conception of the subject who understands, what he terms 'Dasein', and to show how and in what way understanding is an essential feature of Dasein. Heidegger's account has various dimensions and layers. As such, this section is primarily exegetical rather than critical. However, where relevant and possible, I do connect Heidegger's account of understanding with the contemporary literature, and show where there are convergences and divergences. In short, we shall see that Heidegger's own idea of understanding as an ability has its place in this broader investigation of what it means for Dasein *to be*.

As a short disclaimer, Heidegger's terminology and linguistic usage can strike many new readers as at best imaginative, and at worst, barely intelligible. Part of the difficulty is his habit of using philosophical and everyday terms in highly idiosyncratic ways and his penchant for neologism. Throughout this chapter, I have done my best to render his ideas palatable to a

general audience. What I hope to show is that Heideggerian analyses of understanding not only deserve wider recognition within the current debates on understanding, but also have important contributions to make to those debates.

The central idea running through Heidegger's philosophy is that of 'Being'. As such, Heidegger's concern with the subject of understanding, and in turn, the question of what understanding is and what it means to understand, needs to be set within the context of his stated aim in *Being and Time*, which he announces early on as 'reawakening the question of the sense of Being' (*BT*: 19). What does he mean by this?

Being 'is', as Heidegger makes clear, the *intelligibility* of what is (*BT*: 25). For Heidegger, being concerns how we make sense of things; how things in the world can be presented to us as intelligible. To get a grip on what Heidegger means by being, we need to contrast it with another central term. For Heidegger, being is always the being *of* some entity or entities. An 'entity' is Heidegger's term of art for anything and everything that exists: 'Everything we talk about, everything we have in view, everything towards which we comport ourselves in any way' (*BT*: 26). More precisely then, because being concerns intelligibility, then the being of an entity is 'that which determines entities as entities, that on the basis of which entities are already understood' (*BT*: 25-26). The issue for Heidegger then is to determine what is involved in understanding entities *as* entities – that is, what it takes to make sense of them in terms of their being; or rather, in what sense entities are intelligible to us and understood as such. For Heidegger, to consider the 'question of the meaning of being' is thus to consider, as Adrian Moore recently put it, 'the truly remarkable and singular fact that sense is made of anything, and to try and make sense of that' (Moore 2012: 472).

The use of 'as' in the 'entities *as* entities' formulation is meant to capture the fact that in interacting with entities and things in the world, we always interact and comport ourselves toward them in a particular way. I encounter everyday entities as particular things, for example, '*as* a chair', '*as* a pen', '*as* a hammer'. But for Heidegger, the 'as' also indicates a contrast. For example, Heidegger holds that animals and babies do not comport themselves towards entities *as* entities.[6] This is not to deny that animals and babies can obviously interact with entities in all sorts of ways, but that Heidegger thinks that such forms of interaction do not amount to the same kind of interaction that he takes 'Dasein' (and by 'Dasein' he roughly means us as fully-fledged adult human beings) to enjoy. The difference, in Heidegger's view, is the possession of what he calls *ontological understanding*; animals and babies can interact

[6] Heidegger is adamant that *only* human beings (and not even all of them) understand being and thereby interact with entities as entities, but nothing I say in this chapter depends upon that claim being true.

with entities, but, according to Heidegger, they do not possess an understanding of the being of entities.

What does it mean to understand the being of an entity? To understand the being of an entity (including oneself) is to understand the possible ways that entity can be – and importantly, cannot be. Thus, Heidegger characterises ontological understanding in terms of a 'projection of entities onto their possibilities' (*BT*: 182-185). Ontological understanding involves an appreciation of what it is possible for an entity to be and what is not possible for it to be. What this suggests is that there are certain *standards* that we hold entities to in our interaction with them; we have a sense of what a particular entity is or is not allowed to do. It is this 'holding to standards' that offers a rough approximation for what Heidegger means by having an 'understanding of the being of an entity' and suggests that ontological understanding has a *normative* flavour. Let us consider some examples to illustrate Heidegger's point.

To be a hammer is to be able, when wielded in the right manner, to drive in nails (among other things). To understand something 'as a hammer' is to understand it within a field of possibilities that a good working hammer is supposed to be able to capable of. What those possibilities are have been developed through history by human agents engaging in the practice of hammering (and more generally, carpentry). If one were to pick up a hammer that subsequently wilted like a flower, our recognition that this thing is not actually a hammer (perhaps a gimmick or a work of art) is possible because we understand the being of entities such as hammers. In the case of the trick 'hammer', it flouts the *functional standards* that makes hammers what they are, and in this sense, *could not* be a real hammer. In cases like this, one would not have to have a particularly refined understanding to recognise that the fake hammer was not a real hammer, because it flouted quite a general functional standard for what it means for something to be a hammer – it must be hard enough to transfer enough force to other objects. To be a real or actual hammer is in this sense to accord with certain ontological – in this case, functional – standards. It can be used to drive in nails, but it cannot – *qua* hammer – be used to write with (without being used as something quite different), since that is not its intended function. Therefore, to understand the being of an entity like a hammer is to have some grasp of what is functionally possible or impossible for something to be a hammer.

To take another example, consider chess. To play and interact with pieces in a game of chess, players must have at least some minimal sense of what is possible or not for chess pieces to do within the context of a game. That is, to make sense of a chess piece, say, a rook, *as* a rook, players must have some sense of what one can do with a rook and which moves are legitimate or illegitimate for it; in other words, players must have some grasp of the rules of chess. Without the game of chess, understood as an activity prescribed by certain rules and

standards, there would be no such things as chess pieces. In that sense, such rules constitute chess phenomena by providing standards to which what a piece is and what a player does with it are accountable. For example, John Haugeland writes:

> *What* a rook *is*, for instance […] is nothing other than being a piece in a chess game that moves as a rook does (along ranks and files and in castling), and such a piece "*actually exists*" just as long as it fills that role in some ongoing game…What is *possible* in chess is, needless to say, just what the rules enable and allow. Accordingly, the *sense* of the being of chess phenomena – that in terms of which they themselves […] are *intelligible* – is governedness by the rules of chess. (Haugeland 2013: 61)

The rules of chess, including not just the rules governing how pieces are moved, but also rules governing the perceptions and actions of players, are what make chess phenomena (from pieces like rooks to moves like castling to statuses like being in check) intelligible as such; the *being* of chess phenomena, we could say, is their place within such a system of constitutive rules or norms. The rules of chess lay out the field of possibilities for chess phenomena, and understanding within this context amounts to having a sense of what is possible for chess phenomena in accord with the constitutive standards of chess.

To take a more generic example: to be an ordinary, perceptible thing is to behave in more or less stable and predictable ways (Schear 2013: 171). The 'laws' of ordinary, medium-sized things lay out the field of possibilities for being such things. For example, should some object – a chair – pop in and out of existence before one's eyes, something would be awry, and would be recognised as such by any competent perceiver.[7] In this case, as Schear frames it, the disappearing object 'would be flouting the standards of substantial independence and persistence that hold for objects, and so *could not* be any such thing' (ibid: 171). The general idea then is that even mundane, perceptible objects with which we might not have had any previous interaction or experience of, are understood in the light of certain ontological standards.

Finally, to anticipate a later part of the discussion, Heidegger thinks that *something like* this ontological understanding even extends to scientific understanding: in this sense, to have an understanding of the being of entities and phenomena in scientific domains is to have some sense of what is possible or not possible for those entities. Paradigmatically, such ontological standards might be articulated as scientific laws governing a domain, as in physics, but need not be. By specifying what is possible and impossible for physical entities, for instance, the laws of physics specify both what it takes to be a physical entity (rather than not) and what it takes to be *this* sort of physical entity (rather than another). In such cases, laws and other

[7] Although in the first instance, it is more likely that the perceiver would think that something must be wrong with herself, for example that she is hallucinating.

scientific standards, circumscribe the field of possibilities for entities within respective scientific domains.[8]

The basic idea then is that ontological understanding is a matter of appreciating what is possible and not possible for the entities with which we interact in accord with certain standards. To repeat, 'entity' for Heidegger covers anything and everything that exists, and so should not be thought of simply in terms of physical, perceptible objects. For Heidegger, entities refer to anything we can encounter and make sense of: not only people, planets, languages, nations, games, or plants, but also abstract entities such as properties, relations, events, or numbers. To understand the being of any particular entity – to understand that entity *as* an entity – is to have an understanding of what it means to be that particular kind of entity in terms of what is possible or not possible for it. Entities must 'live up' to standards in order to count as being – in order to count as intelligible – and such standards are ontological standards that concern what it is for entities *to be* as opposed to *not* being.

Three further points are necessary before we can move on. First, Heidegger repeatedly stresses that being is *not itself an entity*. This is what he refers to as the *ontological difference*. The basic idea here is that being itself cannot be another kind of entity as such, because if it were, focusing on that entity would only postpone the question of what it means to be; in other words, we would then already be working with an understanding of the being of *that* entity and the target of our investigation would be presupposed, which would lead to a regress. As such, Heidegger distinguishes ontological understanding from *ontic understanding*. The former refers to our understanding of being, whereas the latter refers to our understanding of particular entities ('I understand that such and such is the case' or 'I understand that *x* has these properties'). For Heidegger, ontic understanding always presupposes ontological understanding. It is because we have some sense of what is possible or not possible for an entity to be that we can form specific judgements about entities in the first place. For this reason, while comparisons with the rules of chess or scientific laws are instructive, we should be careful not to directly equate such things, i.e. entities, with being. As such, they are useful analogies rather than direct cases of what Heidegger means by 'being'.

Second, Heidegger thinks that while we all have ontological understanding, it is for the most part vague and unarticulated. Most of us typically only have a vague sense of what is possible or not possible for the entities that we encounter. Some – experts, specialists, trained practitioners – may have a more refined or developed sense: a carpenter, for example, might have a more developed sense of the kinds of possibilities available to hammers than I do, and

[8] The comparison with the physical sciences, such as physics and chemistry, as a means for providing a gloss on Heidegger's characterisation of ontological understanding is one repeatedly made by Haugeland (e.g., 2007, 2013). For a critical discussion of this interpretive strategy, see Withy (2017).

a biologist will have a much more refined understanding of coping with laboratory equipment than a layperson. Sometimes, this vague, pre-theoretical sense of understanding can come to the fore and receives explicit treatment. Thomas Kuhn's notion of scientific revolutions offers quite a useful example of this. Revolutions in scientific thought – evolutionary theory, quantum mechanics, relativity – not only alter the internal dynamics of certain scientific fields, they also fundamentally shift scientists' understanding of what it means for something to *be* an animal or what it is to *be* material (Blatter 2008: 19).[9] Heidegger thus holds that even if our ontological understanding is vague or unarticulated, it is operative in all our dealings with the world, and in that sense, functions as a condition for the possibility of making sense of things at all. Ontology, as Heidegger sees it, is primarily an attempt to articulate our understanding of being.[10] As we shall see in more detail in the next subsection, it is this pre-reflective ability to appreciate and to grasp the distinction between the being and not-being of entities that is essential to human understanding.

Third, Heidegger is an ontological pluralist. This means that there are not just different kinds of entity, but also different ways to be.[11] People, hammers, animals, and mere things all are, but they *are* differently, and Heidegger thinks that we at least implicitly understand and cope with these differences. Heidegger argued that there are different modes or ways of being: different ways in which entities can be intelligible *as* entities. For example, Heidegger argued that at least one mode of being refers to mere things in the world, things and objects devoid of any practical significance. Such things are *present-at-hand*, what Heidegger calls *Vorhanden*, and their way of being is *Vorhandensein* (presence-at-hand, occurrence). Yet, he thought that in our everyday lives for the most part we encounter and deal with things not simply as present-at-hand, but as *available for use*: knives, forks, pens, paper, computers, books, measuring tools, and so on, are entities that are understood in relation to our projects and practices, and in that sense, are not merely present-at-hand. Heidegger refers to this second category as *Zuhandensein* (ready-to-hand, availability). A third category of being is *us*: he thinks that we – as Dasein – are neither merely present-at-hand things in the world, nor are we things available for use; rather he thinks that what makes us distinctive is our ontological understanding, our ability to encounter things as intelligible.[12] Finally, each different way of

[9] Heidegger anticipates these Kuhnian ideas, arguing that the 'real movement of the sciences' (*BT*: 9) occurs in periods of profound conceptual change.

[10] Thus, Heidegger often refers to our vague, unarticulated understanding as 'pre-ontological'.

[11] For a reading of Heidegger's ontological pluralism that connects it with views in contemporary analytic metaphysics, see McDaniel (2013).

[12] Which is not to say that we cannot or do not understand each other as simply occurrent or available – in fact, we often do. The political and social sciences, Heidegger thinks, are paradigmatic ways of understanding us and our interactions with each in terms of availability (*BT*: 413). Similarly, he thinks one of the reasons the Western philosophical tradition has been led astray

being has two aspects, what Heidegger calls their 'that-being' and 'what-being', corresponding to the traditional metaphysical distinction between existence and essence. For example, *what* an item of equipment is is its being specifically located within a larger equipmental context and the social practices that context makes possible (e.g., we understand what the carpenter's hammer is in relation to nails, wooden boards, the general workshop, and so on); *that* there is such an item depends upon its availability or unavailability for the task at hand. In this respect, the that-being of available things does not merely concern existing or not existing, but has a normative dimension. There are, that is, 'deficient modes' of equipmental unavailability: a broken or missing hammer is still a hammer, but in an inappropriate way, and its intelligibility as such derives from understanding its properly functioning availability (*BT*: 102-104).

Ontological understanding – an understanding of the being of entities – is central to Heidegger's account of understanding in general. It should not be thought of as just more knowledge about the world; instead, for Heidegger, it is condition for the possibility of making sense of things in the first place, prior to any claims about the world that might constitute propositional knowledge. Understanding, Heidegger says, 'is not an acquaintance derived from knowledge…but a primordial existential kind of being, which more than anything else, makes such knowledge possible' (*BT*: 161). We must already have some implicit grasp of the intelligibility of things – an understanding of what is possible or not possible for them – before we can form more explicit beliefs about them or entertain whether claims we make about them are warranted or justified. Furthermore, our ontological understanding has a normative character: we understand the being of entities in terms of certain ontological standards and norms. In this respect, 'understanding of being' simply names our orientation towards these standards and norms (Crowell 2012: 240). As I shall suggest in the concluding section, this presents an interesting challenge to contemporary accounts of understanding within epistemology and philosophy of science, which often treat understanding as simply 'super-knowledge' (Lipton 1991) or as a special type of propositional attitude. If a Heideggerian conception of ontological understanding has some plausibility, then it suggests that contemporary accounts have at the very least missed an important *type* of understanding. However, as I have suggested, to appreciate the broader significance of Heidegger's account of ontological understanding, we need to situate it in relation to his conception of the subject who understands, to which we shall now turn.

---

in understanding what it means to be a subject is because it has treated us as just another occurrent thing in the world.

4 Dasein: The Subject Who Understands

Stephen Grimm (2011) and Adam Toon (2015) distinguish three different questions that we might ask about understanding. To introduce these questions, it is helpful to consider parallels with more traditional analyses of knowledge. According to these analysis, S knows that p if and only if p is true, S believes that p, and S is justified in believing that p. On this view, the *object* of knowledge is a true proposition, the relevant *psychological attitude* is belief, and the *normative requirement* is justification. In the formulation 'S understands P', where P is some natural phenomenon, we might ask the same questions. In this case, what is the *object* of understanding? Is it a set of natural propositions about P, or some model or abstract structure? We might inquire into the *psychology* of understanding, and in this case, we might ask whether understanding amounts to believing a set of propositions about P, accepting claims about P, or whether it involves some additional psychological state or ability (which is where appeals to 'grasping' come in). The third issue, the *normativity* requirement, has been the primary focus of epistemologists, who consider whether understanding is subject to the same normative constraints as knowledge. For example, does it require truth?

Toon (2015: 3861) considers a fourth question we might add to Grimm's list. This is the question of the *subject of understanding*. Here the question is *who* or *what* understands the phenomena? As it stands within the contemporary literature, there has been very little engagement with this fourth question. In addition, although various accounts have been put forward that have proposed conditions for what it means for something to be understood, the question of what it means to be a subject who understands and what that involves is rarely asked. To the extent that the subject of understanding is involved, it is often seen in a negative light, as we saw with Hempel, and more recently Trout. Even de Regt's more developed accounts of scientific understanding (e.g., de Regt 2009, 2013), while recognising the pragmatic and subject-involving nature of understanding, typically divert attention away from this issue and focus instead on articulating criteria for understanding a phenomenon. In short, the question of the subject of understanding has been almost entirely neglected within the contemporary literature on understanding. However, as I shall argue in this subsection, that question is of essential importance to Heidegger's account of understanding. To show this, I will focus on three key aspects of Heidegger's account of Dasein: its mode of being (4.1), 'being-in-the-world' (4.2), and the critical of self-understanding (4.3).

4.1 The being of Dasein

Heidegger's term for the kind of entity we are is 'Dasein'. The term 'Dasein' is what Heidegger calls a *formally indicative* term. 'Formal indication' is the use of a term or terms to pick out an entity in a way that circumvents familiar preconceptions of what and how it is. Part of the reason that Heidegger uses 'Dasein' to refer to us, and not a more familiar term like 'subject' is that he takes such terms to come with a lot of unwanted philosophical baggage. Heidegger uses the term 'Dasein' precisely because he wants us to approach and think about the subject of understanding without philosophical preconceptions about what that amounts to or involves.[13] One key preconception that he thinks retains a certain hold over us is a *Cartesian* one, which sees the subject as a mind detached from the world and which sees our understanding as dependent on an 'inner space of meaning' or a 'cabinet of consciousness' (*BT*: 89). As suggested at the start of this section, Heidegger's resistance to the Cartesian approach to understanding is not so much because it rests on an inadequate psychology, as Wittgenstein holds, but rather because it entails a certain conception of the subject as some special kind of thing, a *res*. Now we have a more detailed sense of Heidegger's ontological pluralism, the basic thrust of Heidegger's argument against this becomes clearer. The problem with Cartesianism, Heidegger thinks, is that it commits a category mistake: it treats us as having the mode of being of presence-at-hand (*Vorhandensein*) rather than recognising that what it means to *be* a subject of understanding is not the same as to simply be another kind of present-at-hand thing in the world. Thus, Heidegger's resistance to Cartesianism operates at the level of ontology: the Cartesian conception of understanding presupposes an ontologically mistaken view of what it means to be a subject of understanding.

If Heidegger is right to say that we, as Dasein, have a way of being that is distinctive from the way of being of present-at-hand objects or available equipment, what is the nature of our way of being? We have already seen one way in which Heidegger fleshes out this idea, as Schear puts it:

> To be a Dasein […] is to be a sense-maker, to understand being. Dasein's understanding of being (*Seinsverständnis*) is a capacity or an ability, a *Können*. This, however, is not one ability among others that Dasein might happen peculiarly to have. For unlike being able to tie one's shoes or being able to speak (say) English, understanding being is an ability that is definitive of Dasein as the kind of entity that it is. Hence Heidegger's characterization of this fundamental ability not as an ability one "has" but as an "ability-to-be" (*Seinkönnen*) […] (Schear 2013: 171)

[13] Throughout this thesis, I will nevertheless refer to the 'subject of understanding', but the reader should bear in mind that it is intended in terms of a Heideggerian interpretation of what being a subject is.

What distinguishes Dasein from other entities, then, is that Dasein is the kind of entity that has an understanding of being and as such is able to make sense of entities *as* entities (*BT*: 34). Heidegger's term for Dasein's mode of being is *existence*. Thus, in *Being and Time*, Heidegger refers to at least three modes of being: *presence-at-hand*, *readiness-to-hand*, and *existence*. As might be expected, Heidegger's use of the term 'existence' runs counter to more familiar philosophical usage. It does not refer to the fact that human beings simply are. Rather, Dasein's way of being as existence means for it to be constantly directed towards certain possible ways in which it can be. This will probably sound obscure, so I will try to clarify it further in the following two points.

First, Heidegger thinks that Dasein's being - that is, how Dasein makes sense of and understands itself – is always at issue for it (*BT*: 67). He thinks that Dasein is the kind of entity that is fundamentally concerned with how it lives its life and what form that life takes. This concern often manifests itself in times of importance and change within our personal lives: do I continue to try and be an academic philosopher? Do I take on this new job offer? Who am I trying to be? For Heidegger, such choices and possibilities have meaning and significance insofar as we understand ourselves in terms of them, and thus, whether and how we take up and follow one set of possibilities over another is something that is *at issue* for us. The second point is that the being which is an issue for Dasein, 'is in each case *mine*' (*BT*: 67; emphasis added). In one sense, this simply draws out an implication of the first point: for any entity that chooses to live in a particular way, that entity makes that existential possibility its own. As Steven Crowell argues, these aspects of Dasein's existence are imbued with normative significance. The possible way in which Dasein can be functions as a kind of standard to which it holds itself accountable: '…to "be" a father is for what it means to be a father to be at issue for me in trying to be one: I do not merely do certain things but commit myself to the possibility of failure. That is, for *me* being a father is a normative status' (Crowell 2012: 215). In thus understanding myself in a certain way - paradigmatically in the form of socially prescribed, practical identities – what I say and do is accountable to what it is that I am trying to be.

These two aspects of Dasein's mode of being as existence distinguish it from the way of being of other entities. For example, tables, chairs, and rocks cannot make sense of themselves, or understand themselves in terms of their own being. As such, they do not, in Heidegger's technical sense, *exist*. Whereas such objects have properties which go towards classifying them as what they are, Heidegger argues that Dasein has – or rather is – its possibilities (*BT*: 183). What distinguishes Dasein is its ability to make sense of and understand itself in terms of the possible ways in which it can be. This is why Heidegger

characterises Dasein's ontological understanding as not just any contingent cognitive faculty or ability, but one which is essential as an 'ability-to-be' (*Seinkönnen*).[14]

On Heidegger's view, then, understanding is essential to what it means to be Dasein. Dasein is defined as the kind of entity that has an understanding of being, which means to be able to make sense of entities *as* entities. In addition, it is the kind of entity for whom the way in which it continues its life can be at issue for it. As we shall see, for Heidegger, this means that any specific case of understanding is not just a matter of having the right kind of beliefs or knowledge about the world, nor just about possessing a certain set of psychological or cognitive abilities. Instead, for Heidegger, understanding presupposes *self-understanding*. Fundamentally, Heidegger thinks that making sense of human understanding cannot be done without reference to what it means to be the subject of understanding.

4.2 Being-in-the-world

Heidegger thinks that we all possess ontological understanding, but that for the most part it is vague, unarticulated, and 'pre-theoretical'. The idea here is that while we all have some sense of what it means for different entities to be and we are able to cope with those differences, we ordinarily do not make that understanding itself the subject of inquiry. It remains implicit in our activities and various dealings with the world. What I want to do in this section is focus more closely on how Heidegger characterises the basic structure of our engagement with things in the world and to develop the sense in which our understanding is a fundamental part of entities like us.

As we have seen, Heidegger distinguishes between three modes of being in *Being and Time*: the present-at-hand (*Vorhanden*), the ready-to-hand (*Zuhanden*), and existence (*Existenz*). Heidegger argues that, for the most part, we ordinarily encounter and understand entities as (what he calls) *equipment*; that is, as being for certain sorts of tasks (cooking, writing, building, and so on). Equipmental entities, as we have seen, have availability as their mode of being. Heidegger's focus on our understanding of equipmental entities is philosophically significant for several reasons, but primarily, he thinks that philosophy has for

[14] Earlier I noted that each different mode of being has two different aspects: a 'what-being' and a 'that-being'. W*hat* an item of equipment is is determined by its location within a larger equipmental context; *that* there is such an item depends upon its availability or unavailability for the task at hand. Applied to Dasein's mode of being, the 'what-being' of Dasein *is* depends on the specific configuration of possibilities that it has taken up in living its life. The 'that-being' ties into Heidegger's later discussion of *authenticity* and *inauthenticity* in Division II of *Being and Time*, which refer to different ways in which Dasein can be. Discussing the significance of these existential aspects for Heidegger's account of understanding lies beyond the focus of this chapter however. For a recent discussion of these themes, see McManus (2015).

too long focused only on our theoretical, detached relations to things, which is characterised in purely cognitive terms. The paradigm of this way of understanding would be something like Descartes' observation of his ball of wax, detached from any immediate practical task, and the complex array of other objects and people within which such tasks are pursued (Mulhall 1996: 41). Such a way of engaging with entities is appropriate for understanding present-at-hand entities, but, Heidegger thinks, fails to do justice to the more pervasive *practical* form of understanding that shows up in our everyday practices and projects. Thus, Heidegger suggests that 'understanding something' is manifest as practical competence: being able to manage something and carry out specific tasks (*BT*: 183). Making sense of entities in terms of their determinate properties and features is certainly one way of making sense of them, but by no means the only way. For example, Heidegger thinks that such an understanding of a hammer does not characterise how we understand things like hammers when we are actually using them:

> The less we just stare at the hammer-thing, and the more we seize hold of it and use it, the more primordial does our relationship to it become, and the more unveiledly is it encountered as that which it is – as equipment. The hammering itself uncovers the specific 'manipulability' of the hammer. The kind of being which equipment possesses - in which it manifests itself in its own right – we call 'readiness-to-hand'. (*BT*: 98)

Thus, at an everyday level, Heidegger takes our ontological understanding to have a pragmatic character. In this respect, his account of our practical competence and abilities bears important similarities to Gilbert Ryle's account of 'knowing how' (Ryle 1946). In addition, the use of available entities, like hammers or other pieces of equipment, has a distinctive phenomenological signature, according to Heidegger. He argues that while engaged in such activity, Dasein has no conscious experience of the items of equipment in use *as independent objects* (i.e., as the bearers of determinate properties that exist independently of the context of action in which the equipmental entity is involved). Instead, our action and understanding of them manifests what Hubert Dreyfus would later call 'skilful/absorbed coping': the ability to engage with one's circumstances by responding to the demands of the situation in a skilful and often effortless manner, without conscious deliberation, reasoning, or planning. Heidegger thus thinks that in using available entities, they take on a certain transparency. In using it, phenomenologically, the tool itself becomes transparent: my focus is not on the tool itself but on the task to be achieved, the job to be done. When writing on my laptop, I do not pay attention to the laptop itself – instead I am directed towards the words being typed on the page. As a particular kind of tool, the laptop itself is phenomenologically transparent. Similarly for the biologist in using a microscope or preparing a sample: if the tools function as they are supposed to, her attention is not on the tools themselves but on the task at hand

and the goal to be achieved. In this sense, the biologist can properly 'see through' the microscope because it does not inhibit the activity, if functioning properly. It is only when something goes wrong that our equipment starts to 'stand out' and require our attention. For example, if my laptop crashes, I become aware of it as a particular kind of entity within a practical context. Crucially, this transparency in use does not mean that there is no awareness in our practical activity or that our behaviour is simply routine or automatic. Rather, what the phenomenology of practical competence and skilful coping suggests is that the awareness that is present (what Heidegger calls *circumspection*) is not a matter of a subject encountering objects before them. As Wheeler summarises, phenomenologically speaking, 'there are no subjects and objects; there is only the experience of the ongoing task' (Wheeler 2017). Thus, our practical understanding does not consist in the thematic representation of an object to a subject but in a 'concernful attending to and noticing what is significant within the practical nexus of one's surroundings' (Rouse 2002: 51).

It is important to be clear, however, about what Heidegger is and is not arguing by focusing on availability, readiness-to-hand and the pragmatic nature of understanding. For one, he is *not* arguing that the philosopher's traditional primacy of the theoretical and focus on the present-at-hand should be replaced by a primacy of the practical and a focus on the ready-to-hand.[15] Heidegger does not take readiness-to-hand to be *metaphysically* prior to presence-at-hand. What he *does* claim is that the traditional focus on the present-at-hand and the contemplative attitude towards things has obscured how the way in which we understand things manifests itself in other sorts of activities beyond the theoretical. He does think there is a certain *phenomenological* priority to the ready-to-hand, because it is how we primarily engage with things in the world. But to reiterate, Heidegger's point is not that *all* understanding is practical, but that the practical character of our understanding has been passed over by the philosophical tradition.[16]

One crucial aspect that has traditionally been obscured by this focus on a theoretical, detached view of things is that available entities are intelligible in reference to a broader, equipmental context in which they are placed. For example, understanding a hammer as available for use means understanding it in relation to other pieces of equipment – nails, wooden boards, the workshop – what Heidegger calls a 'referential totality' (*BT*: 70, 105). In this sense, individual items within this equipmental context 'refer' to each other: the hammer is only intelligible in reference to the rest of the equipmental context. It is precisely because

[15] Although this interpretation of the philosophical significance of Heidegger's work was for a long time the dominant one, as defended by Dreyfus (1991).

[16] For further discussion of this point, see McManus (2012), particularly chapters 3, 4 and 6.

of the contextual nature of our understanding of equipment that leads Heidegger to the claim that,

> Taken strictly, there 'is' no such thing as *an* equipment. To the being of any equipment there always belongs a totality of equipment, in which it can be this equipment that it is. Equipment is essentially 'something in-order-to…'. A totality of equipment is constituted by various ways of the 'in-order-to', such as serviceability, conduciveness, usability, manipulability. (*BT*: 97)

Here, we have a version of what Robert Brandom calls Heidegger's 'strong systematicity condition': that any entity encountered as available is encountered within a holistic network of practical affordances (Brandom 1983: 391-3). Such referential connections are not causal connections, but are characterised in functional terms. Equipmental entities are understood as 'in-order-to' do something, which Heidegger characterises as 'involvements' (*BT*: 115):

> With any such entity as entity there is some involvement. The fact that it has such an involvement is ontologically definitive for the being of such an entity […] with this thing, for instance, which is ready-to-hand, and which we accordingly call a 'hammer', there is an involvement in hammering; with hammering there is an involvement in making something fast; with making something fast, there is an involvement in protecting against bad weather; and this protection 'is' for the sake of providing shelter for Dasein. (*BT*: 116)

So, Heidegger thinks that the kind of practical understanding involved in our dealings with available entities is always in relation to some particular goal. As such, the functional significance of these equipmental references also has a normative structure. The series of 'in-order-to' relations do not simply amount to a circumspective awareness of a particular piece of equipment, but rather indicate that such practical contexts and tasks are normatively contoured by conditions of success or failure. Such conditions structure the sort of things the use of equipment must achieve to count as a successful instance of its relevant activity. As Mark Okrent notes, a practical understanding of hammering involves not simply the understanding 'use this *x* to do *y*', but also a grasp of *when* it is appropriate to use a hammer, the *circumstances* in which it is appropriate to use one, and the appropriate sort of *way* (Okrent 1988: 32-33).

More extensively, Heidegger also holds that the normative conditions are shaped by the kinds of social and institutional roles which are appropriate to them (e.g., what it means to be a carpenter), which in turn make certain actions appropriate and others inappropriate within a given context. Thus, Heidegger famously argues that before being an individual subject, Dasein is first and foremost a socialised entity, which he calls 'the One' (*das Man*). There are two aspects to this. On the one hand, it means that the entities themselves are bound up with norms of appropriate use: for something to be a hammer means that it comes up with certain norms of how to use it appropriately. As such, ready-to-hand entities are understood in terms

of their roles which are determined by the place they hold in our social practices. On the other hand, this normative dimension also fundamentally shapes my own actions, responses and performances, which, in turn, are shaped by the social expectations that are bound up in the various practical identities and roles I take up. For example, being a teacher is bound up with social norms concerning what it is appropriate for teachers to do, what to say, and, Heidegger thinks, how to make sense of things in the world; that is, being a teacher is a particular kind of 'practical identity' which comes with certain socially and culturally determined norms and expectations. Heidegger describes this as our *thrownness*: we are 'thrown' into a particular socio-historical set of practices and we are primarily oriented towards the world in terms of these practices and traditions. The historical, social, and cultural setting for the most part determines the kinds of possibilities in terms of which we understand entities.[17]

The general point that Heidegger is making here is that the way in which I make sense of things in the world is neither dependent simply on some stock of propositional knowledge about things, nor even exhausted by my practical abilities to navigate the world; rather, it belongs to the 'public', the historically and culturally particular social practices of those among whom I find myself (Crowell 2012: 202). The point is that Heidegger holds that 'a necessary condition for the possibility of encountering things as meaningful is the habitual conformity to public norms' (ibid: 202). Thus, while Heidegger would agree with acknowledging the importance of 'knowing-how', his own account of the practical nature of our understanding goes beyond this notion by situating it in relation to this broader practical context, shaped by functional relations, social and institutional roles, and our ongoing projects.

These contextual interrelations that underpin our practical understanding lead into one sense of what Heidegger means by the term 'world'. Crucially, Heidegger characterises Dasein as 'being-in-the-world' (*BT*: §12). This means that as subjects of understanding, we are always situated within a meaningful context where we deal with and encounter other things and people, directed towards particulars goals and purposes which provide conditions of satisfaction for the success or failure of our actions. The practical understanding of Dasein is one way in which Heidegger fleshes out what being-in-the-world means: in this case, to be in the world is to be oriented towards entities in terms of their practical significance and to have an understanding of them in relation to their use, function, and equipmental contexts. But there are other important features that Heidegger thinks are necessary for understanding what it means for Dasein to 'be-in-the-world'.

'Being-in-the-world' is a unitary phenomenon, which means that each part of it is essential to the other and cannot be understood separately from one another. In the first instance, Heidegger is clear that 'being-in' should not be construed in spatial terms (*BT*: 79-80). For

[17] A useful discussion of Heidegger's notion of 'thrownness' can be found in Withy (2011).

Dasein to be in the world is not simply to refer to its spatial location in relation and proximity to other entities in the world. Rather, 'being-in' captures the idea that my own orientation to the world and the way in which the world presents itself to me are structured by patterns of meaning and significance. In Heidegger's usage, 'being-in' means to be familiar with things, which in turns means to know how to cope with them. The character of this 'familiarity' is indicated by the paradigmatic ways in which it occurs:

> Such possible modes of in-being belonging to everydayness include: working on something with something, producing something, cultivating and caring for something, putting something to use, employing something for something, holding something in trust, giving up, letting something get lost, interrogating, discussing, accomplishing, exploring, considering, determining something. (Heidegger 1992: 159)

Heidegger thinks that the common factor to these aspects is *concern*: 'Dasein finds "itself" proximally in *what* it does, uses, expects, avoids – in those things environmentally ready-to-hand with which it is proximally *concerned*' (*BT*: 155). We have a concern and an interest in the world – things *matter* to us – such that our engagement with the world is shaped by the meaningful affordances of the environment in which we are situated.

It is because Heidegger thinks we always find ourselves 'in' the world in this way that it is a mistake to treat the subject of understanding as just another present-at-hand entity. Here, Heidegger's critique of Descartes comes to the fore (*BT*: 128-129). According to Heidegger, Descartes presents the world to us as a collection of present-at-hand entities to be encountered by us subjects. A Cartesian view of human existence treats it as enclosed within a special inner sphere – the mind, ego, consciousness, or the self. Outside this sphere stand external objects – brute, material things situated in measurable space. Heidegger challenges this Cartesian view in many ways, but within the context of his discussion of 'world', the basic critique is this: the Cartesian view is fundamentally at odds with the phenomenology of our lived experience of the world. If the Cartesian view were right, this would imply that Dasein would only be able to engage with things as ready-to-hand (in terms of practical, equipmental significance) by adding 'value-predicates' to this foundational context-independent stratum of present-at-hand entities (*BT*: 132-133). This would suggest that meaning is first cooked up in the inner sphere of the subject and then projected onto an otherwise meaningless world. Heidegger argues that this picture gets the phenomenology all backward: meaning is not simply projected onto external objects. Rather, we always already find ourselves in meaningful contexts. As Dasein we are first and foremost directed towards available, ready-to-hand equipmental entities, which are intelligible and laden with context-dependence significance. Thus, Heidegger writes,

> What we 'first' hear is never noises or complexes of sounds, but the creaking waggon, the motor-cycle. We hear the column on the march, the north wind, the woodpecker tapping, the fire crackling […] It requires a very artificial and complicated frame of mind to 'hear' a 'pure noise'. The fact that motor-cycles and waggons are what we proximally hear is the phenomenal evidence that in every case Dasein, as Being-in-the-world, already dwells alongside what is ready-to-hand within-the-world; it certainly does not dwell proximally alongside 'sensations'; nor would it first have to give shape to the swirl of sensations to provide a springboard from which the subject leaps off and finally arrives at a 'world'. Dasein, as essentially understanding, is proximally alongside what is understood. (*BT*: 207)

As we shall see later, this has implications for Heidegger's approach to scientific understanding. For he thinks that, while scientists may encounter entities in more of a 'detached' sense, treating objects of inquiry as bearers of measurable and determinate properties and features, this is only possible because the scientist is already situated within a broader context – the scientific 'world' – which forms the meaningful backdrop against which some entities can stand out as objects of inquiry. Heidegger's point then is that even a scientific investigation of the world should not be thought of primarily in terms of encountering a brute, material stratum, which we then impose layers of meaning or values upon. Rather, scientific inquiry presupposes (and Heidegger thinks is only possible because of) the establishment of meaningful contexts which direct and determine what is significant or relevant to the aims of inquiry. As such, the Cartesian conception of the subject and the world suffers from what Richard Polt calls 'intellectualist myopia': 'it examines a specialized class of entities within the world, present-at-hand objects, but it misses the world itself, the larger context in which these objects present themselves' (Polt 1999: 58).

To sum up: to be a subject of understanding is to be situated within historically and culturally developed, norm-governed social practices, which are themselves located within the meaningful whole of a 'world', and Dasein is predominantly carried along by the standards, values and concerns of its concrete, social world. Phenomenologically, our understanding primarily takes the form of an implicit, practical competence: a kind of skilful know-how in terms of which we are able to cope with our environment, the entities within it, and the possibilities that pertain to them.

4.3 Dasein's self-understanding

Understanding, for Heidegger, is a fundamental aspect of our being-in, or familiarity with, the world, insofar as without our understanding, Heidegger thinks we would not be able to encounter or engage with things in the world at all. Thus, as we have seen, Heidegger treats understanding as a fundamental *existentiale*: an essential component of what it means to be a subject. We are only able to encounter, interact with and comport ourselves to entities *as* entities because, Heidegger thinks, we understand the being of those entities; that is, because

we can make sense of them in terms of their possibilities. As being-in-the-world, our understanding is primarily manifest in a kind of practical competence with entities: an ability to cope and deal with things in terms of the possibilities appropriate to them in relation to specific contexts and social practices. The preceding subsection set out one central way in which these possibilities are determined: in terms of their practical significance and context-dependence. However, one way that Heidegger significantly departs from other accounts of understanding is in his emphasis on *self-understanding*. Heidegger takes Dasein's self-understanding to be an essential factor in how it makes sense of the world at all: it provides something like a foundation for all the other forms that understanding can take. Heidegger thinks that our practical understanding presupposes that Dasein understands itself in a particular way. It is in self-understanding that the various aspects of Heidegger's account of understanding that we have discussed so far come together.

Recall that our practical understanding of entities always means having a grasp of their place in a network of equipmental significance. I understand the hammer in relation to the nails, wood and workshop, or the laptop in relation to the desk, books, and cup of coffee. Each part of this milieu make sense in terms of its relation to other parts. Moreover, this referential context has a teleological structure: we use particular things 'in-order-to' achieve and carry out some task, which Heidegger characterises as 'involvements'. Such involvements do not have a uniform structure however. Instead, there are various different 'levels'. For example, I am currently working with my laptop (what Heidegger calls a 'with-which'), in the practical context of a library (an 'in-which'), in order to write this chapter (an 'in-order-to'), with the aim of completing my thesis (a 'towards-which'), for the sake of my academic work, which is to say, for the sake of my *being* an academic (a 'for-the-sake-of-which'). The crucial point then for Heidegger is that any practical understanding of things ultimately comes back to some possible way in which Dasein can *be* in the world. For Heidegger, I encounter and make sense of things in the world in relation to a 'for-the-sake-of-which', which is formed on the basis of how I understand myself as existing (in Heidegger's technical sense of existence as the mode of being of Dasein) in a certain way. Thus, how we understand things in the world ultimately refers back to us as the kind of entity that can act 'for the sake of' a particular practical identity. The significance of this is that it means that the intelligible possibilities in terms of which we understand entities are related to Dasein's own projects and possibilities.

The activity of writing this chapter is understood only insofar as I understand myself as a student or academic. Furthermore, the spaces I inhabit, the ways in which I am 'in' the world, are constituted in relation to the goals and purposes that make up what it means to be an academic, in relation to certain socially prescribed and publically available norms. The library, department, study spaces and coffee shops all have a particular significance and salience for me, insofar as they show up as meaningful contextual possibilities towards which my actions

and practices make sense. But more importantly, it is because of this link to my own self-understanding that such possibilities have normative force: the practices of being an academic *matter* to me only because I am trying to *be* an academic. It is because what it means for me to be an academic is something that can be at stake or at issue in what I do – prescribing the terms for which I can succeed or fail – that my understanding of entities in the world have normative significance. Thus, with a slight modification, Steven Crowell argues,

> To try is not merely to act in accord with norms (mechanically, as it were) but to be responsive to the normative, to the possibility of living up to the demands of what it is to be a writer or failing to do so. A monkey could perhaps try to write but it could not try to *be* a writer; could not, in Heidegger's terms, understand itself as a writer, act in the light of writerly norms (Crowell 2012: 29)

Crowell's point here is that in trying to live and act in accord with some practical identity, Dasein must be understood as the kind of entity that not only acts in accord with certain norms (the norms of being a writer), but responds to norms *as norms*. Dasein is the kind of entity for whom norms have normative force because who Dasein is and how it continues its life can be at issue for it. Being responsive to norms is thus not just an incidental feature of what it means to be a subject of understanding, according to Heidegger, but is an essential part of it.

Thus, Heidegger argues that to understand entities in terms of their possibilities ultimately presupposes the way in which Dasein understands itself in terms of its possibilities. According to Heidegger, the way in which we understand things in the world and the way in which we understand ourselves are two sides of the same coin. In understanding things in the world by way of their place in these broader contexts of significance, Dasein not only understands entities in terms of what is possible for them (at least as far as the social and practical norms governing those activities are constitutive of those entities), but also in so doing, understands itself in relation to those possibilities. Dasein, in its ability-to-be-in-the-world always has at its root self-understanding. This understanding is neither something Dasein has 'in mind' so to speak, nor is it simply a species of cognition; instead, it is manifest in its very ability to deal with things. This is ultimately why in Heidegger's view, one cannot make sense of human understanding without making sense of what it means to be a subject who understands in the first place.

In this section I have focused on Heidegger's account of what it means to be a subject of understanding. Heidegger's term for the subject who understands is 'Dasein'. He defines Dasein as the kind of entity who has an understanding of the being of entities, the way in which they can be or cannot be. Furthermore, we saw that Heidegger argues that Dasein's mode of being is *existence*, where existence is defined as the ability of Dasein to understand *itself* in terms of its own possible ways to be, such that, the way in which Dasein can be (how

and in what way it continues its life) can be at issue for it. As being-in-the-world, Dasein's understanding, according to Heidegger, is for the most part manifest in our everyday, practical engagement with things in the world. We encounter entities primarily as available or ready-to-hand, in the form of equipment. The possible ways in which equipmental entities can be is circumscribed by the practical contexts in which we encounter them, and the concrete, norm-governed world in which we are situated. To be 'in the world' means to be located in a particular socio-historical, cultural setting, in terms of which we encounter things in terms of their familiarity, significance and concern. The significance of entities is underpinned by their place in a totality of involvements, which ultimately leads back to some practical identity that Dasein understands itself in terms of. This self-understanding delineates the way in which Dasein makes sense of itself in terms of the kinds of possibilities it chooses for itself, paradigmatically in the form of practical identities (e.g., a writer, a teacher, an academic, a father). This self-understanding, I have suggested points to the way in which Dasein is the kind of entity that can be responsive to norms, such that it not merely acts in accord with them, but also acts in the light of them. Because of this responsiveness, the way in which Dasein chooses to exist *matters* to it, such that it can succeed or fail at whatever way of life it has chosen.

5 Implications for Scientific Understanding

Heidegger's account of science in *Being and Time*, while very suggestive, is underdeveloped and involves a number of difficulties.[18] Moreover, there is little to suggest the extent to which Heidegger engaged with the philosophy of science of his day, with the exception of a few polemical remarks.[19] Nevertheless, there are three features of Heidegger's account of scientific understanding that I wish to highlight in this section: (i) its practical character; (ii) its prospective nature; and (iii) its normative dimension.

The central claim of Heidegger's existential conception of science is that science should be understood as something Dasein does, as a specific way in which it comports itself to the world (*BT*: 408). Heidegger explicitly contrasts this *existential* conception of science to more

[18] As an area of research, Heidegger's philosophy of science has received less attention than other aspects of his philosophy and their relation to other branches of philosophy. Nevertheless, recent work by Patricia Glazebrook has done much to rectify this, in particular Glazebrook (2000, 2012). For a brief overview of some of the elements and problems with Heidegger's philosophy of science, see Rouse (2005).

[19] As Rouse notes, Heidegger's development of an existential conception of science would have primarily been critically directed towards the neo-Kantians and Husserl's transcendental phenomenological approach. For an excellent discussion of Heidegger's relation to both neo-Kantian philosophy and the Vienna Circle, see Friedman (2000).

traditionally epistemological conceptions that identify science as 'the totality established through an interconnection of true propositions' (*BT*: 32), and do so through 'the kind of "logic" which limps along after the sciences in order to discover its "method"' (*BT*: 30). In effect, Heidegger is criticising traditional approaches within the philosophy of science which are primarily concerned with the retrospective evaluation and justification of scientific results and its methodological norms. For Heidegger this approach to science rests of the assumption that, on the one hand, the relevant scientific research has been completed, and on the other hand, that what is to be evaluated can be made fully explicit. As Rouse (1981, 2005) emphasises, Heidegger is not concerned with this reconstructive project, but rather, insofar as science is understood as something Dasein *does*, the sciences must be considered *prospectively*:

> Research demands that we project possible directions in outline, in advance of their actually having being carried out. To make such understanding fully explicit would be to have *already* carried out the research activities which the understanding is projecting. A fully explicit preconception would not tell us what would count as a solution to a problem; it would be the solution. (Rouse 1981: 274)

For Heidegger, it is the retrospective conception of scientific knowledge that is precisely what science actively seeks to surpass. For Heidegger, the aim of an existential conception of science is to thematise this prospective, futural-orientation of science, in which science, as a way in which Dasein directs itself to the world, always 'presses ahead' into unexplored possibilities.[20] Existentially conceived, Heidegger argues that the sciences must be approached as 'a way of existence and thus as a mode of being-in-the-world, which discovers or discloses either entities or being' (*BT*: 408). That is to say, the sciences manifest specific ways in which Dasein can exist in the world, a particular body of practices through which entities are intelligibly discovered as entities, which in turn presupposes an understanding of their being.

On a Heideggerian reading, the contexts in which scientific research is undertaken, the laboratory, clinic or field site, are not a collection of extant objects with various properties and relations, but instead constitute 'worlds': normatively structured contexts of significance. The practical understanding provides the background of possibilities against which the relevant skills, techniques and abilities of scientists enable them to successfully interact with, discover and manipulate the equipment with which they are concerned. On this Heideggerian reading, scientific instruments are not simply theory-driven 'tools' to which we put 'definite

[20] This prospective character of scientific research has been more recently discussed in Hans-Jörg Rheinberger's account of experimental systems (Rheinberger 1997).

questions'.[21] Instead, they are determined by our practical understanding of the kinds of possibilities that are available to them as equipment and a sensitivity to the normative conditions which determine their accuracy, reliability, repeatability and relevance for inquiry. Such understanding, as Heidegger emphasised, need not involve an explicit representation of the 'work to be produced', but a non-thematic grasp of the possibilities available to equipmental entities, and a sense of how to act within these equipmental contexts (*BT*: 185).

Heidegger emphasises the practical dimension of scientific understanding by arguing against conceptions of scientific or theoretical inquiry which see it as involving detachment from our instrumental and equipmental concerns. If we characterise scientific inquiry as a shift from everyday circumspective concern to theoretical exploration, then theoretical understanding would consist primarily in the 'disappearance of praxis' or a 'privation' and 'absence of praxis' (*BT*: 409). But, according to Heidegger, scientific research is underpinned by bodies of research practices:

> Reading off the measurements which result from an experiment often requires a complicated 'technical' set-up for the experimental design. Observation with a microscope is dependent upon the production of 'preparations'. Archaeological excavation, which precedes any interpretation of the 'findings', demands manipulations of the grossest kind […] even in the 'most abstract' way of working out problems and establishing what has been obtained, one manipulates equipment for writing, for example. (*BT*: 409)

Heidegger's point here is not to reduce all scientific activity to a kind of practical know-how in opposition to theoretical inquiry. His concern is rather to emphasise that our practical orientation to the world is not a mere 'add-on' to scientific inquiry but a constitutive element of the manner in which the sciences understand and make sense of entities within their disciplinary fields. The idea that science consists simply in pure theoretical inquiry has of course long been challenged by philosophers of science in the Anglo-American tradition, which is illustrated in the work on scientific experimentation, instrumentation and the material dimension of scientific practice.

However, although Heidegger does suggestively point to the dimension of practical understanding in science, his remarks on this matter are quite limited. In one respect, it is arguable that Heidegger did not want to characterise scientific research as simply a technical or pragmatic instrumentalism. Science, as a way in which Dasein comports itself to the world, does not simply discover entities through experimental and technological expertise. Instead, for Heidegger, scientific research involves disclosing the being of entities; that is, scientific

[21] In the classical formulation of Karl Popper, 'the theoretician puts certain definite questions to the experimenter, and the latter, by his experiments, tries to elicit a decisive answer to these questions, and to no others' (Popper 2002: 89).

inquiry takes place with its own understanding of the being of entities. Thus, although it is possible to make sense of scientific understanding along a practical dimension, it is by no means exhaustive for the ways in which scientific understanding makes sense of entities. Instead, any scientific domain embodies a *regional ontology*, which is a particular understanding of the being of entities within that domain. As a regional ontology, a science's understanding of being determines what counts *as* an entity in its domain; it provides the conditions under which an entity can be intelligibly disclosed within a given domain of inquiry. Scientific understanding thus involves a projection of the being of the entities encountered in scientific research. As Rouse argues, Heidegger's ontological pluralism, in which there are many different ways in which entities can *be*, stressed this point:

> What a regional ontology thus projects is what is *possible* for any entity within that domain, and hence also what is *impossible* for it. Empiricists since Hume have been suspicious of modal attributions within the sciences, because empirical experience seems only to show what is actually the case, and not what *must* be so or *can* be so. That suspicion reflects what Heidegger thought to be a characteristic limitation of any retrospective, non-existential conception of scientific understanding. Scientific understanding does not provide an encyclopaedic summary of past experience, but instead projects how to encounter and deal with entities in a domain. (Rouse 2015b: 95)

Heidegger's account of science thus emphasises the centrality of understanding in guiding scientific research, in determining what it means to count as an entity within a particular domain, and articulates the conditions for what is possible or not possible for entities in that domain. For Heidegger, the sciences are practices of research rather than bodies of knowledge extracted from those practices. His account of the ontological nature of understanding implies that such practices depend upon a prior understanding of the being of entities within the research domain, in the form of a regional ontology, which governs scientific methodology and conceptualisation.

6 Concluding Remarks

The aim of this chapter has been to orient the basic direction of this thesis. To do this, I have drawn upon Heidegger's conception of human understanding, which emphasises both its pragmatic and normative character. In contrast to contemporary discussions of understanding, Heidegger provides an account of what it means to be a subject of understanding. Heidegger's account of Dasein as being-in-the-world reveals our understanding to be an essential factor in how we engage with and deal with things in the world. One finds in Heidegger's characterisation of Dasein's understanding certain themes that will recur throughout the thesis. First, we understand entities in terms of what it is possible for them to do and to be, and such

possibilities are prescribed by ontological standards and norms of intelligibility. Second, our understanding is primarily manifest in terms of practical competence, abilities and know-how, which in turn is situated within norm-governed social practices. Third, at the root of Heidegger's account of the subject who understands is the idea that, fundamentally, Dasein is the kind of entity who can not only act in accord with norms, but also in the light of them, such that it is capable of responding to norms *as norms*. In what follows, I shall argue that understanding does not just mean being responsive to norms, but also taking *responsibility* for them: in understanding things in the light of norms, I undertake a commitment and responsibility to uphold those norms, to not tolerate violations of them, and to question them when they fail to connect to how the world is. Therefore, Heidegger's focus on the subject of understanding represents a significant contribution to the contemporary debate on understanding, which has so far not addressed the question of what it means to be a subject who understands. How we make sense of the subject who understands has important implications for what it means to understand as such, and the assumptions we make at that level will inform how our accounts of understanding will be developed.

Chapter Two

Understanding as Inferential Ability

1 Introduction

Chapter One proposed a conception of the subject of understanding based on Heidegger's phenomenological ontology. On this conception, to be a subject of understanding is first and foremost to be situated in a meaningful and intelligible world of norm-governed, social practices, and to be capable of responding to norms *as* norms. On this reading, understanding is a practical, normative capacity to deal appropriately with one's surroundings and grasp things in terms of their possibilities. These possibilities are ontological, in the sense that understanding something means to understand it in terms of the possible ways it can or cannot *be*.

In this chapter, I build on this Heideggerian conception of understanding by combining it with Robert Brandom's 'normative inferentialism'.[1] Brandom's inferentialism is concerned with *conceptual understanding*: it is primarily an account of what is involved in understanding conceptual and semantic content. Like Heidegger, Brandom emphasises the pragmatic character of conceptual understanding and its connection to norm-governed, social practices. On this view, conceptual understanding is an inferential ability governed by rules of material inference. One's understanding is determined not by the actual inferences one can draw, but instead by the appropriateness and legitimacy of those inferences in the light of social-normative inferential rules. Therefore, the broader aim of this chapter is to show how conceptually articulated understanding depends upon the normative pragmatics of our social practices.

The chapter will proceed as follows. Section Two motivates the turn to conceptual understanding and characterising it in terms of inferential ability. In Section Three, I introduce the explanatory strategy of Brandom's inferentialism, which can be understood as an attempt to ground an inferentialist semantics in a normative pragmatics. Section Four outlines Brandom's normative pragmatics in terms of the paradigmatic normative statuses that

[1] Brandom's inferentialism is systematically presented in *Making It Explicit* (Brandom 1994), hereafter referred to as *MIE*. The same themes are developed in *Articulating Reasons* (Brandom 2000), which I shall refer to as *AR*.

constitute a social, linguistic practice and how these statuses relate to the 'game of giving and asking for reasons'. Section Five turns to the inferentialist semantics in more detail, focusing on the notion of material inference, while Section Six extends this account to incorporate a two-dimensional structure of conceptual normativity understanding. Section Seven concludes by identifying key points of convergence and divergence between Brandom's normative inferentialism and the current debates on the notion of understanding.

2 Motivating Inferentialism

Inferentialism is primarily concerned with what it means to have *conceptual* understanding: it offers an account of what is involved in our grasp of conceptual and semantic content. The nature of conceptual understanding has not typically been addressed in the current literature on understanding. As we saw in the Introduction, the contemporary debates have largely focused on determining the relationship between understanding and knowledge on the one hand, and between understanding and explanation on the other. The focus of these concerns has been distilled into a distinction between two types of understanding: *objectual* understanding and *explanatory* understanding.[2] Objectual understanding is understanding a thing, a set of related things, or a subject matter and is typically characterised as involving the ability to connect various items of information together in what Kvanvig calls 'coherence-making relationships' (Kvanvig 2003: 192). Explanatory understanding, by contrast, is having an explanation for why something is the case. However, there is currently no agreed conception of explanatory understanding. Does it, for example, only require knowledge of particular explanations (Khalifa 2012; Strevens 2013)? Or does it involve further kinds of abilities that are constitutive of that understanding, such as counterfactual reasoning (Grimm 2011; Hills 2015) or the ability to derive explanations from broader theoretical frameworks (de Regt 2009)?

In what sense does an account of conceptual understanding pertain to the types of understanding discussed in contemporary debates? First, it is important to recognise that conceptual understanding is presupposed by both objectual and explanatory understanding. We acquire objectual understanding on the basis of our ability to discern appropriate relations and connections between our concepts and the world. Similarly, explanatory understanding is

[2] Explanatory understanding is discussed by Grimm (2010), De Regt (2009a, 2013), Khalifa (2012), Strevens (2013) and Alison Hills (2015) among others. Objectual understanding has been discussed by Baumberger (2011), Elgin (2007), Kvanvig (2003), Khalifa (2013), and Kelp (2015). Of course, various other types of understanding have been distinguished (e.g., propositional, interrogative, moral, aesthetic, and so on), but these two types of understanding have been the primary focus of discussion in epistemology and philosophy of science.

dependent on the specific kinds of concepts that are embedded within explanations: be they causal, nomological, mechanistic, functional, and so forth. As such, investigating the nature of conceptual understanding is an important prerequisite for 'higher-level' forms of understanding such as found in objectual and explanatory understanding, where a basic grasp of conceptual and semantic content is already in place.

Second, an inferentialist approach to conceptual understanding holds that inferential ability is constitutive of understanding. This represents a significant convergence with contemporary discussions on understanding, which similarly take inferential ability and performance as an important criterion of understanding. For example, several authors have argued that inferential ability is a plausible candidate for fleshing out the metaphorical notion of 'grasping' that is often invoked in claims of understanding. According to Stephen Grimm (2006, 2011), grasping an explanation requires the ability to infer correct descriptions of similar, possibly counterfactual cases. For Grimm, we understand why something is the case by anticipating how changes in some of the explanatory details leads to changes in others.[3] Similarly, Alison Hills (2015) has suggested that the distinction between merely believing or knowing an explanation and understanding it depends on the degree of 'cognitive control' one has over the variables involved in an explanation. She argues that one understands why *p* when (i) one can follow an explanation for why *p*, (ii) explain why *p* in one's own words, (iii) draw the conclusion that *p* from *q*, (iv) make analogical inferences from *q'* to the relevantly similar *p'*, (v) use *p* to explain other related propositions, and finally, (vi) use *p'* analogically to explain related propositions, *q'* (ibid: 663). In turn, Daniel Wilkenfeld (2013) and Mark Newman (2012) have drawn on recent developments in cognitive science to argue that understanding involves the possession of mental representations or mental models of the object of understanding, which the agent is able to manipulate in ways that 'enable efficacious inferences' (Wilkenfeld 2013: 1004). For Wilkenfeld, a necessary condition of understanding is that an agent possesses a mental representation of the object of understanding. Understanding *qua* grasping, on this account, consists in the cognitive act of manipulating internal, mental representations of the objects and vehicles of understanding in order to infer to new cases. Similarly, Newman proposes that,

> […] when we attribute to someone the state of understanding a scientific explanation for some phenomenon, *p*, we are not taking them to have merely literal knowledge of what is going on that generates *p*, nor do we think that they must be able to derive *p* themselves from only initial

[3] Grimm is explicitly drawing on James Woodward's (2003) manipulationist account of scientific explanation. According to Woodward, explanation consist of tracing or exhibiting functional dependency relations between variables. For Woodward, these explanatory relationships provide understanding by facilitating answers to 'what-if-things-had-been-different' questions concerning the consequences of counterfactual or hypothetical changes in the values of the *explanans* variable.

> conditions and laws. Rather, we are attributing to them the mental state of having made appropriate inferences on each part of the explanatory story that ends with *p* as its conclusion. (Newman 2012: 15)

Although I disagree with Wilkenfeld's and Newman's characterisation of an agent's inferential ability in terms of mental representations and models, their accounts nevertheless serve to highlight that inferential ability is an important factor in our capacity to understand. It captures the sense that understanding involves not merely the possession of beliefs, knowledge or concepts but also our grasp of the connections between them, our ability to use and manipulate that knowledge, and ultimately to reason about the world. A robust theory of inferential ability, which I argue we find in Brandom, would thus represent a significant contribution to a theory of understanding.

Third, and perhaps most significantly for an account of scientific understanding, an inferentialist approach to conceptual understanding resists characterising scientific knowledge and understanding in fundamentally representational terms. Scientific understanding is irreducibly discursive in character and is permeated with language: what the sciences allow one to *say* about the world is thus an important component of any account of scientific understanding (Rouse 2002: 185). It is often presumed that a representationalist approach to semantics is the only viable candidate that can account for this. Yet representationalism has always carried with it the crucial problem of how our knowledge and understanding of the world is linked to the world, if there is no access to anything except via another representation. The standard philosophical response to this worry has been to build our knowledge on what is 'given' in sensory experience, or to look for 'privileged representations' (Rorty 1979) that would ground that knowledge: immediate ideas, concepts or logical forms. Such representations would then somehow automatically grant knowledge or understanding. In turn, this quest for privileged representations has also informed the cognitive sciences, where hypothetical entities such as concepts, symbolic structures, mental models, prototypes and schemes are ascribed to our minds in order to explain our cognitive capabilities (Knuuttila 2011: 264).[4] But this reduction of scientific understanding to a set of privileged representations about the world fails to do justice to the rich complexity and diversity of practices of scientific research. What the development of coherent research practices enables one to do and say is significantly more extensive and complex than the philosophical focus on representational accuracy or explanatory insight would suggest. This is not to suggest that representational

[4] Knuuttila further notes (2011: 264) that recent discussions of scientific representation, particularly in the context of the epistemic value of scientific models, are one variant of this general representationalist predicament, insofar as they have typically centred on whether or not there is a privileged set of structural relations that underpins scientific representation, in the form of some isomorphism between a model and its target.

import and accuracy or the use of diverse kinds of representational media within scientific practices are irrelevant to scientific understanding; far from it. Instead, the point is to resist taking representation as the basic explanatory primitive for making sense of scientific understanding. An inferentialist approach to conceptually articulated scientific understanding offers, as I shall argue in Chapters Four and Five, a rich set of conceptual resources and tools for such an account.

3 Normative Inferentialism

The basic claim of inferentialism is that meaning (i.e. conceptual content) should not be analysed in terms of *reference* but in terms of *inference*. The fact that what we say and do, that a statement means something is a fact that we should try to understand not in terms of the relation between a statement and a state of affairs, but in terms of inferential relations between a statement and other statements. However, although inferentialism is first and foremost a semantic theory, concerned with how what we say has meaning and content, Brandom proposes to ground semantics in pragmatics. Pragmatics is concerned with what one is *doing* when using linguistic expressions, which means Brandom analyses conceptual content and understanding in terms of what one is doing when using particular terms and expressions. This order of explanation thus builds upon the later Wittgenstein, in which meaning is analysed in terms of use. To clarify Brandom's particular version of inferentialism, we need to contrast it with two other approaches to semantic content: (i) *semantic representationalism*, and (ii) *causal inferentialism*.

First, Brandom opposes what he takes to be the orthodox representationalist approach to semantics. Representationalism attempts to account for the semantic content of linguistic expressions by invoking various forms of word-object or word-world relations. As a very basic characterisation, the representationalist picture holds that words are contentful/meaningful insofar as they represent, or stand for, something (on the basis of suitably worked out truth- and/or reference-conditions). Whole sentences are then explained as being composed out of the meanings of the words they comprise, which then can be used to explain which inferences are valid. Furthermore, a representationalist semantics is often, but not necessarily, coupled with the idea that 'the mind's representational capacities precede, cause and explain linguistic meaning and what one is doing in using language' (Maher 2012: 78).

Semantic representationalism can be thought of as the conjunction of two assumptions (Wanderer 2008: 96). The first claims the explanatory priority of semantics over pragmatics in providing a philosophical account of language. The assumption here is that we first need to

secure our semantic theory before moving on to understand our use of expressions. The second assumption holds that semantic notions like truth and reference are explanatory primitives. The representationalist approach thus adheres to a particular order of explanation with regard to language: we (1) begin with an account of truth and reference as the basic semantic concepts; (2) proceed from there to explain meaning and inference in terms of truth and reference; before finally moving to (3) the pragmatic level, an account of what we are capable of doing with our semantics in place. This implies a division of labour, such that semantics and pragmatics can be treated as conceptually independent from one another. While philosophers might agree in principle that a full-blown theory of language will eventually need both, in practice, they are approached with relative autonomy.

The basic move of Brandom's inferentialism is to reverse this order of explanation. Like the representationalist approach, it can be presented as proceeding in three stages. The first stage of Brandom's alternative begins in an account of our linguistic practice: an account of what we are doing when using expressions. The specific kind of use or activity that Brandom primarily focuses on is the activity of *asserting* which I shall discuss in greater detail below.[5] With his account of linguistic practice in place, the next stage develops our understanding of semantic content in terms of its inferential role. The aim of the third stage is to offer an account of the representational dimension of our language in terms of the previous two stages. Like the representationalist order of explanation, each stage thus builds on the previous one. Wanderer (2008: 97) schematises Brandom's general strategy as the following:

> B1 An account of linguistic practice, incorporating the speech act of asserting
>
> B2 An account of the semantic content of linguistic items in terms of the inferential role of such linguistic items. The concept of inferential role is explained in terms of such linguistic items being caught up in the linguistic practice outlined in B1.
>
> B3 An account of the representational dimension of semantic content, including the use of terms such as "true" and "refers", in terms of the notion of inferential role outlined in B2.

Thus, Brandom's inferentialism challenges and rejects both assumptions that motivate semantic representationalism: (1) the idea that semantics precedes pragmatics in the order of

[5] Brandom writes, 'The explanatory strategy pursued here is to begin with an account of social practices, identify the particular structure they must exhibit in order to qualify as specifically *linguistic* practice, and then consider what different sorts of semantic content those practices can confer on states, performances and expressions caught up in them in suitable ways' (*MIE*: xiii).

explanation, and (2) that truth and reference are semantic primitives. Both rejections are undoubtedly controversial and Brandom is by no means the first to argue for the priority of pragmatics over semantics. However, it is relevant for the following discussion that we recognise the constraints Brandom's commitment to this order of explanation place on his account. The key point is that for each stage to do explanatory work - that is, to explain the notions and concepts introduced in later stages – Brandom cannot appeal to concepts that occur in later stages on pain of begging the question. For example, as we shall see, Brandom cannot and does not appeal to notions of meaning and content to explain linguistic practice. Similarly, he cannot and does not appeal to notions of truth to explain valid inference. In the next section, I shall only sketch out the basic contours of Brandom's first two stages, B1 and B2, which provide the main ingredients for my approach to the concept of understanding. This does not mean that I take the third stage B3 to be irrelevant, but that it would need more space to work out adequately.

Setting out these starting points for Brandom's account helps to clarify the inferentialist approach to conceptual understanding. If Brandom's order of explanation for conceptual content can be treated as an analogue for an account of understanding, then it would suggest that we treat understanding first in terms of what we do in our practices, rather than making sense of understanding primarily in representationalist terms. Another implication is that the representational capacity of understanding should be treated as dependent on this more basic pragmatic understanding, a point that we saw in Heidegger's conception of understanding. To reiterate the point, the basic pragmatist commitment underlying Brandom's inferentialism is to explain knowing *that* in terms of knowing *how* (see MIE: 135-6).

The second contrast between Brandom's version of inferentialism and others concerns the notion of inference itself. An inferentialist semantics holds that semantic content is a function of inferential role: content is conferred on our expressions in virtue of its role in inference and reasoning, rather than in terms of its origin in experience (as an empiricist might hold). Stated as such, Brandom's inferentialism bears resemblance to what is known as *inferential role semantics* or more generally, *conceptual role semantics*,[6] which claims, in the words of Ned Block, that the 'meaning of a representation is the role of that representation in the cognitive life of an agent, e.g., in perception, thought and decision-making' (Block 1998).

It is necessary, however, to distinguish Brandom's approach from these other varieties of inferentialism to avoid confusing them. Although both stress the idea that meaning and content be considered as a function of inferential role, the crucial difference between them lies in how they make sense of this role. Consider Boghossian's characterisation of inferential role semantics in the following:

[6] For example, as discussed in Fodor and Lepore (1992) and Boghossian (1993).

> Let's suppose that we think in a language of thought and that there are causal facts of the following form: the appearance in *O*'s belief box of a sentence S_1 has a tendency to cause the appearance therein of a sentence S_2 but not S_3. Ignoring many complications, we may describe this sort of fact as consisting in *O*'s disposition to *infer* from S_1 to S_2, but not to S_3. Let's call the totality of the inferences to which a sentence is capable of contributing, its *total inferential role*. A subsentential constituent's total inferential role can then be defined accordingly, as consisting in the contribution it makes to the total inferential role of the sentences in which it appears. (Boghossian 1993: 73)

On Boghossian's construal, the inferential role that confers content on an expression is characterised first and foremost in *causal* or *dispositional* terms. The focus is accordingly on what inferences the subject actually makes or is disposed to make. By contrast, Brandom's inferentialism primarily concentrates on the inferential *rules* and *norms* that govern the inferences made by the subject, rather than the actual inferences a subject makes. That is, Brandom's approach shifts attention to the normative proprieties that govern our inferences, rather than being concerned with the actual inferences themselves. As such, the relevant distinction here is between *causal inferentialism* and Brandom's *normative inferentialism* (Peregrin 2014: 8).

This marks a significant contrast with the kind of account of inferential ability proposed by Wilkenfeld or Newman, discussed in the previous section. While the ability to infer is something that one does on the basis of certain cognitive abilities, inferential rules are not themselves states or events in the causal order. Brandom's normative variety of inferentialism is committed to an account of the content of our thoughts and expressions in terms of what we *ought* to do if what we say and think is to have determinate semantic content. Normative inferentialism holds that the ability to grasp conceptual content, and thus to count as understanding, is a matter of the *proprieties* and *norms* of inference, rather than in terms of the dispositions of subjects to infer one thing rather than another. For Brandom, 'the *content* of any concept is the conceptual content that it is only in virtue of playing a specific normative role or having a certain normative status within the logical space of reasons' (O'Shea 2010: 462). In the same sense that something only counts as a chess piece in virtue of being subject to the rules of chess, for the normative inferentialist, meaning and content are only understood in relation to such inferential rules. Meaning, content and understanding are thus dependent on being situated within a rule-governed, inferentially articulated space. Thus, Brandom's account of inferential content should be understood in the light of Saul Kripke's interpretation of Wittgenstein's rule-following problem, which concludes that an adequate account of meaning and content must not simply describe actual behaviour or dispositions to behave. Rather, 'the relation of meaning and intention to future action is *normative* and not *descriptive*' (Kripke 1981: 37).

To sum up: in this section I have introduced two distinctions to clarify the nature of Brandom's approach to conceptual understanding. The first distinction was between representationalism, which understands conceptual content in terms of word-world relations, and inferentialism, which argues that conceptual content acquires meaning through its use in reasoning. The second distinction was between causal inferentialism, which explicates inferential ability in terms of the inferences the subject actually makes or is disposed to make, and normative inferentialism, which focuses on the rules and norms that govern the appropriateness and legitimacy of those inferences. For Brandom, conceptual understanding is less about the particular cognitive abilities one may or may not possess, and more about the way in which we are responsive to norms. Conceptual understanding, in this respect, is more about being able to appreciate the distinction between correct or incorrect ways in which we take the world to be and to be responsive to the norms governing our inferential practices.

4 Normative Pragmatics

The aim of the first stage of Brandom's account of meaning and content is to work out a model of linguistic practice that does not appeal to semantic notions. For Brandom, a linguistic practice is a type of social practice. I shall discuss how norms and practices are mutually inter-definable in the following chapter. However, it is necessary to anticipate some of the content here in order to understand Brandom's normative pragmatics.

Norm-governed social practices depend, for Brandom, on the normative attitudes taken by practitioners towards themselves. Such normative attitudes are to be understood primarily as ways of treating one another in terms of the correctness or appropriateness of what we say and do, and in that respect, are expressions of the more general attitude of holding one another to account. Because of these normative attitudes, a social practice can be thought of as a constellation of performances that can have the normative status of being correct or incorrect, appropriate or inappropriate, significant or insignificant, and so forth. Social practices are thus conceived on an 'I-Thou' model of sociality and are identified based on the mutual accountability of agents to another (*MIE*: 39, 62). A linguistic practice is a specific kind of social practice built out of certain kinds of performance which are taken to have a normative status. Accordingly, we can say that all linguistic practices are social practices, but not all social practices are linguistic. A central claim of Brandom's account is that a social practice qualifies as linguistic if it includes the activity of *asserting* as a dominant kind of performance within that practice.

The privileging of the act of asserting here requires some qualification. It is a matter of some debate whether Brandom is justified in prioritising the act of assertion as a putatively

necessary and sufficient condition for something to qualify as a linguistic practice. Brandom's reasons for this seem to be that other speech acts are ultimately dependent for their intelligibility on the ability to assert something. For example: 'Saying "Shut the door!" counts as an order only in the context of a practice that includes judgements, and therefore assertions, that the door is shut or that it is not shut' (*MIE*: 172). Roughly, the idea seems to be that other speech acts would not be possible if one were not able to assert that such and such is the case. This assumption, I think, is problematic, and has been rightfully criticised. Rebecca Kukla and Mark Lance, for instance, argue that Brandom succumbs to what they call the 'declarative fallacy', an erroneous and pervasive assumption 'that the structure of *declarative assertions* is the *privileged or sole* dimension of language' (Kukla and Lance 2008: 11). The problem with this fallacy, they argue, is that it unnecessarily blinds us to important pragmatic differences between our speech acts, instead treating them as secondary or derivative upon a fully fleshed out account of assertion.[7]

Although Brandom does not provide a specific argument for why the act of asserting should be taken as basic, his commitment to a general rationalist outlook does go some way in explaining it (if not justifying it) (Wanderer 2008: 20). Brandom draws from Kant the insight that what marks us apart from other creatures is that we can take *responsibility* for what we say and do, in the sense that we can be held responsible for giving *reasons* for what we say and do, and moreover, are capable of responding to reasons *as* reasons, seeing them as being binding upon us. Brandom follows Kant (and Frege) in holding that the minimal unit of cognition that one can be held responsible for is a *judgement*, expressed in an assertion, rather than the individual subsentential parts – terms, concepts, predicates - of which it is composed (*MIE*: 79-80). Consequently, Brandom's Kantianism commits him to taking the act of asserting to be to central because it is the primary linguistic unit which can stand both in need of reasons and as providing reasons for other claims.

Asserting something, for Brandom, is making a move in what Sellars called the 'game of giving and asking for reasons' (*MIE*: 167; *AR*: 189). In making an assertion, I simultaneously undertake a kind of justificatory responsibility – to respond to appropriate linguistic challenges, to provide evidence and reasons for my assertion – and to offer to others an inference license – to endorse, that is, 'claims that can be appropriately inferred from the assertion in question and to take up a stance of challenge or incompatibility to a range of

[7] Brandom's focus on assertion is also arguably at odds with Heidegger's focus on the more basic way in which our understanding is first and foremost a matter of practical comportment to the world. However, I overlook this possible difficulty for the purposes of my more general strategy of identifying certain themes concerning the nature of understanding at work in both thinkers.

contrary assertions' (Lance 2008: 410).[8] Saying or thinking that something is a door amounts to (at least implicitly) asserting that it is a door, and that means undertaking a number of inferential commitments, commitments to say or think a lot of other things as well: that it can be open or closed, that it is a way to leave or enter rooms, that it is attached by hinges, and so on. Brandom thus privileges assertion because he sees it as the primary form in which one can offer a reason for something. This connects with the idea of inferring as the basis for treating one claim as a reason for another:

> […] *assertions* are fundamentally fodder for *inferences*. Uttering a sentence with assertional force or significance is putting it forward *as* a potential reason […] Assertions are essentially *fit* to be reasons. The function of assertion is making sentences available for use as premises in inferences. (*MIE*: 168)

On Brandom's construal, asserting is thus fundamentally normative and intersubjective, because 'asserting is understood in terms of what it *allows* and *requires* [agents] to do in relation to one another' (Maher 2012: 67). Framing the activity of asserting in these normative terms underpins why Brandom's account should be understood as rooted in a *normative pragmatics*: Brandom's strategy is to show how the pragmatic norms governing linguistic and discursive practice can account for the content of our conceptual understanding. At the root of this strategy are the two basic normative statuses of the game of giving and asking for reasons: *commitments* and *entitlements*.

To assert that *p* is to undertake a commitment to *p* and to take up the role of one to whom challenges to *p* may be directed. In asserting *p*, an agent thus undertakes a responsibility to provide reasons for *p* if challenged by other members of the discursive, linguistic community. Commitments come in two distinct forms: (1) commitments that are *avowed* or *acknowledged* and (2) commitments that are *attributed* by others. When an agent asserts that *p*, she is avowing or acknowledging that *p*; it is what the agent takes herself to be committed to. In doing so, other members of the linguistic practice also take that as evidence for *attributing* the commitment that *p* to the agent. Both forms of commitment are thus in Brandom's idiom not understood as *de facto* psychological states, but rather as public, normative statuses.

The notion of 'commitment' is Brandom's substitute for the notion of 'belief'. In making this substitution, the idea here is to avoid thinking of belief primarily in psychological terms.

[8] The Kantian connection between understanding and responsibility is an important aspect of the way in which understanding is fundamentally connected to our place within a normatively structured 'space of reasons'. If we fail to take responsibility for what we say or do in the form of responding to, giving, and asking for reasons, then our claim to understanding is undermined. In that respect, we could see responsibility as potential necessary condition for understanding. I discuss this possibility further in the conclusion.

In that respect, while Brandom would see belief/commitment as a condition on understanding, this should not be seen as a narrowly psychological attitude to propositional content, but rather as the adoption of a public, normative status which presupposes the capacity of the agent to be both responsive to the normative attitudes embodied within social practices, but also to take responsibility for the content of their claims about the world.[9] As with the 'feeling' or 'sense' of understanding, this is not to deny that we do often take psychological attitudes to what we say and do, but rather to argue that the public, normative dimension of our understanding retains a kind of conceptual priority.

The complexity of discursive, linguistic practice reveals itself in these committive consequences because what else the agent understands herself to be committed to in virtue of being committed to *p* may differ from the consequences that other agents take a commitment to *p* to entail. These differences may reflect, for example, differences in one's training, education or understanding. In undertaking a commitment to *p*, the agent may acknowledge certain consequences of this commitment which may be different from the ones attributed to her by other members of the practice. It is possible, and in most cases quite likely, that one will not acknowledge all such consequences even though, in ideal circumstances, one should. The normative dimension of this is that in undertaking a commitment, one is 'binding oneself beyond that which one currently recognizes as the committive consequences of the undertaking' (Wanderer 2008: 43). The differences between those commitments acknowledged and those commitments attributed by others thus marks a difference in perspective between members of the linguistic practice. On the one hand, there is what the agent takes herself to be committed to, and on the other hand, there is what others take the agent to be committed to (Maher 2012: 68). Thus, commitments are normative statuses, attitudes to conceptual content that one acknowledges, attributes, and undertakes in an essentially social context.

This brings us to the second normative status, that of *entitlement*. We have seen that when an agent asserts that *p*, she undertakes a commitment and responsibility to provide reasons for *p* if challenged by other members of the practice. Brandom argues that it is a basic feature of linguistic practice that the agent who asserts that *p* has a *default* entitlement to *p*. That is, other agents must assume that she is entitled to *p* unless they have good reasons not to think so.

[9] It is a matter of some debate whether belief should be seen as a condition on understanding, in the same sense as traditional analyses of knowledge as non-Gettiered, justified true belief, and how it should be characterised. See, for example, Grimm (2006, 2010), Kvanvig (2003), Hills (2015), and Pritchard (2010). Others have argued that the relevant condition should be a broader notion such as acceptance or commitment (e.g., Elgin 2004; Dellsen 2016). In that respect, Brandom would sit closer to the latter camp, with the notion of commitment understood first and foremost as a normative status adopted within a social practice.

Moreover, as there are consequences to our commitments, there are also consequences of entitlement. In asserting that *p*, I am not only entitled to *p*, but also its 'permissive consequences'. Entitlement, in this sense, is inherited: it is conferred over the range of consequential commitments accrued to an agent. For example, if I am entitled to the claim that "This ball is red", I am also entitled to the claim that "This ball is coloured".[10] Once again, there may be differences between what the agent takes herself to be entitled to and what other agents take her to be entitled to. Furthermore, for Brandom our commitments are not only typically implicit in practice: our grasp of the consequences of those commitments is by no means always transparent to agents. The consequences of our commitments often extend beyond what we are capable of articulating, and as such, we frequently may have inconsistent sets of commitments. For Brandom, working out what one is entitled to and should endorse is part of the ongoing nature of discursive inquiry. As such, the idea that one is entitled to certain commitments by default does not mean that one's entitlements to a commitment can go unchallenged. Indeed, Brandom argues that, on this model of linguistic practice, the justification of what is asserted has a *default and challenge structure* (*MIE*: 176). On this model, agents have entitlement to their commitments by default, but can lose that entitlement through a challenge if, for example, an agent is unable to answer that challenge.[11] In doing so, we are constantly refining and reshaping our inferential commitments.[12]

Agents can respond to challenges either by (1) simply disavowing their commitment; (2) offering reasons for their entitlement to that commitment; or (3) by showing how other permissive consequences do follow if entitlement to *p* is preserved. If A's entitlement to *p* is maintained, then other agents are entitled to assert that *p*, unless it is incompatible with any of their respective commitments. In this way, agents inherit entitlement to commitments from each other. Thus, avowing a commitment can also be taken as the issuing of a 're-avowal license' to other agents (Maher 2012: 70). Again, the complexity of linguistic practice

[10] This is an example of what Brandom, following Sellars (1953), calls 'material inference', according to which inferences are licensed (supported or warranted) by the content of the concepts that figure in the propositions which are inferentially connected. Whether an inference is licensed or not depends on the content of the concepts the inferences are made about. This contrasts with formal inference, according to which an inference is valid based on their logical form, and thus has traditionally been the type of inference studied by classical logic. I discuss the notion of material inference in detail in the following section.

[11] Brandom's 'default and challenge' model of justification has been discussed with reference to traditional epistemological conceptions of justification in Williams (2001).

[12] For example, I argue in Chapter Four that the construction, development, and analysis of models in scientific practice is one way in which scientists refine their conceptually articulated scientific understanding. Models are treated as assemblages of conceptual commitments, and model-based reasoning is understood as working out what the inferential consequences of such commitments are.

involves recognising that what counts as a commitment, a permissive consequence, or an entitlement for an agent, depends on the different contexts, perspectives and normative positions in the game of giving and asking for reasons.

This serves to introduce the essential elements of Brandom's normative pragmatics. In this account, linguistic practice is modelled on a game (of giving and asking for reasons) in which agents undertake and attribute commitments. Avowing, acknowledging and undertaking a commitment alters the 'score' of both the agent herself and other participants. Exactly how a particular avowal affects the agent's score will depend on her pre-existing commitments and entitlements. For Brandom, to participate in linguistic and discursive practice is to take on the role of a 'deontic scorekeeper' (*MIE*: 142), a role which involves keeping track of the various normative statuses that they and other agents hold, and monitoring the altering scores of different 'players' within that practice. Brandom thus foregrounds the normativity of discursive practices, through the implicit normative statuses of participants within that practice. What matters here is that participants in social practices take normative attitudes to one another and that through their mutual interaction, they continually (and implicitly) hold each other to account by monitoring the scores of other agents in the discursive practice.

5 Inferentialist Semantics

Brandom's account of linguistic, social practice explicates the act of asserting in terms of normative pragmatics. This provides the groundwork for his inferentialist semantics, which shall be the focus of this section. My aim is to show how these two aspects – a normative pragmatics and an inferentialist semantics - fit together. In the following section, I will bring the preceding discussion together to present an inferentialist account of conceptual understanding.

An inferentialist approach to conceptual understanding holds that meaning and conceptual content is conferred on what we say and do on the basis of its inferential role in our discursive practices. The use of an expression acquires conceptual content from the appropriateness of when and how it is used. Meaning and content are inferentially articulated and accordingly inferentialism embraces *semantic holism*. Since meaning and conceptual content are a matter of inferential relations and their inferential significance depends on their relations to other concepts, 'one must have many concepts in order to have any' (*MIE*: 89). Moreover, since the inferential connections one endorses are in turn affected by the kinds of background commitments and entitlements one has, no two speakers will attach the same significance to an utterance or expression. This is usually interpreted as a problem for the semantic holist,

insofar as it would appear to render impossible communication between two agents.[13] For instance, if, because of his very different collateral commitments, Rutherford meant something quite different by 'electron' than I do, it would seem to be impossible to disagree with him about whether electrons have fixed positions and orbits, since what I take to be the content of the concept 'electron' is radically different from his (Brandom 2007: 663). Concerning the issue of the cognitive process of science, this issue would seem especially pressing. Brandom's response to this objection is to appeal to inferential norms. The claim is that communication is possible, not because of shared or conventional meanings, beliefs and content between agents, but because the inferences that an agent makes are understood in terms of their accountability to the norms governing those inferences. As such, speakers count as using one and the same sentence-type if and only if they are held accountable to the same inferential norms for their use of tokens of that sentence. As Brandom illustrates,

> The norms I am binding myself to by using the term 'molybdenum' – what actually follows from or is incompatible with the applicability of the concept – need not change as my views about molybdenum and its inferential surround change. And you and I may be bound by just the same public linguistic and conceptual norms in the vicinity in spite of the fact that we are disposed to make different claims and inferential moves. (AR: 29)

Thus, Brandom wants to give up the atomistic idea that we associate meaning and content with our concepts one by one, instead holding that concepts are contentful in virtue of their standing in inferential relations to other concepts and intersubjective norms (Brandom 2007: 662). The inferentialist holds that meaning and content is a function of inferential roles, the kind of inferences that a sentence is caught up in. How then is the meaning of a sentence determined? Brandom distinguishes three varieties of inferentialism which offer different responses to this question: *weak inferentialism*, *strong inferentialism*, and *hyperinferentialism.*

Weak inferentialism holds only that 'the inferential connections among sentences are *necessary* for them to have the content that they do' (ibid: 656). The idea here is relatively uncontroversial. It states that inferential connections form a necessary part of the semantic content of an assertion, such that if a sentence were involved in (at least some) different inferential connections, it would mean something different. Brandom conjectures that most semantic theorists, even representationalists, would agree to this weak form of inferentialism.

Strong inferentialism holds that this kind of 'inferential articulation' (i.e. the kinds of inferential connections that sentences enter into) is not only a necessary, but also a *sufficient* condition for meaningfulness. This contrasts with *hyperinferentialism*. The difference between strong inferentialism and hyperinferentialism concerns the scope of these inferential

[13] See, for example, Fodor and Lepore (2001).

connections. Hyperinferentialism holds that inferential connections among sentences are necessary and sufficient for meaning, but takes the sufficiency of inferential connections to be *narrow* in scope, whereas the strong inferentialist takes them to be *broad* in scope. The narrowness of hyperinferentialism is exemplified insofar as it models inferential connections *only* in terms of introduction and elimination rules for logical vocabulary, such as logical connectives.

Let me explain. Introduction and elimination rules can be used as a way of defining the meaning of logical connectives in ways that do not contain the connective being defined. For the connective '&', an introduction rule would state that anyone committed to *p* and committed to *q* is committed to *p* & *q*; the elimination rule would state that anyone who is committed to *p* & *q* is committed to *p* and committed to *q*. The connective is thus defined in terms of both the inferentially sufficient conditions for use of sentences containing the connective, and in terms of the inferentially necessary consequences of the use of sentences containing the connective. Brandom, however, takes this narrowness of scope to be untenable for an account of semantic content in natural language, which importantly should include observational and empirical terms, that is, terms that do not simply have meaning in virtue of logical introduction and elimination rules, but also in terms of their non-inferential applicability to things in the world (*AR*: 28). The main difference between the hyperinferentialist and the strong inferentialist is that the hyperinferentialist takes these introduction and elimination rules to be *all there is* to inferential articulation, whereas the strong inferentialist relaxes this idea, accommodating reliable perception and action within the range of what is inferentially significant.

Strong inferentialism therefore aspires to be a *via media* between weak inferentialism and hyperinferentialism, and is the version of inferentialism that Brandom defends. It holds that inferential connections are both necessary and sufficient for meaning, but it claims that these connections are broad in their scope of application. The broadness of scope advocated by the strong inferentialist can be summed up in the three ways in which it departs from the hyperinferentialist model (Brandom 2007: 657-8): (1) that not all good inferences are formally valid ones, but can also be *materially* good ones; (2) that the type of materially good inferences constitutive of the meaning of a claim also include material incompatibilities; and (3) that inferential relations between *non*inferential circumstances of appropriate application and *non*inferential appropriate consequences of application are also taken into account. I shall discuss each of these in turn.

5.1 Material Inference

Central to Brandom's inferentialist semantics and account of conceptual understanding is the notion of material inference, following Sellars (1953). According to material accounts of inference, inferences are licensed (supported or warranted) by the content of the concepts that are involved in the inference. Thus, whether an inference is licensed depends on the content of the concepts the inferences are made about. This contrasts with formal inference, which holds that all valid inferences are licensed by the logical form of the propositions that are inferentially connected with each other. Thus, the inference from

(a) this substance turns litmus paper red
    to
(b) this substance is an acid

is a materially good inference, but formally invalid. From the perspective of standard logic, the inference from (a) to (b) can only become formally valid with the addition of another premise:

(a1) any substance that turns litmus paper red is an acid

This standard view thus treats the inference from (a) to (b) as enthymematic, that is, as requiring supplementation with the quantified conditional claim (a1). However, Brandom rejects the formalist presumption that material inference should be treated as enthymematic. As such, he argues that material inference has conceptual priority over formal inference. There are three reasons why Brandom takes the notion of material inference, rather than formal inference, to be central to an inferentialist semantics.

First, Brandom takes it that most inferences in natural language take the material form. For example, 'It is raining, so the streets will be wet'; 'There is lightning, so there will be thunder'; 'I am releasing this piece of chalk, so it will fall' are all on this construal materially good inferences, despite being formally invalid. Crucially, Brandom thinks that we normally take it that such material inferences are warranted without requiring an additional premise. The warrant stems from the content of the concepts involved, rather than from the form they take. For Brandom, the ubiquity of material inference in everyday practice respects Wittgenstein's insight that meaning is implicit in practice (meaning is use) and that the competent use of language is possible without being able to lay out explicit, formal rules of language, formulated in some previously learned metalanguage (Sellars 1954).

Second, Brandom endorses Sellars' argument for the indispensability of material inference and its central role in conceptual understanding. The basic claim in this argument is that material inferences are *counterfactual supporting*, a feature which Sellars argued could not be captured in formal rules of inference, but which nevertheless plays an important role not only in scientific language but in any language. Sellars contends that for this added premise (a1, above: 'any substance that turns litmus paper red is an acid') to have any normative force over the validity of the inference – for it to have the force of an inferential rule - it must not only generalise over actual cases of substances that turn litmus paper red and acids, but also possible and counterfactual cases too. That is, 'Any substance that turns litmus paper red is an acid' would have to be read as a subjunctive conditional: 'If anything *were* a substance that turns litmus paper red, it *would be* an acid'. More precisely, it would have to be read as a rule saying that for any *x*, '*x* is an acid' may be inferred from '*x* turns litmus paper red'. Sellars argues, however, that to read subjunctive conditionals in this way is in effect to re-introduce material rules of inference into the argument, because 'this rule ['*x* is an acid' may be inferred from '*x* turns litmus paper red'] is not in any obvious way a specification of a purely logical rule of inference' (Sellars 1953: 323). Expressing the added premise in this way, Sellars holds, is rather to codify in subjunctive conditionals the commitments that are already and implicitly involved in our understanding of material inferences. *Contra* the formalist, it follows that the correctness of the move from the premise to the conclusion cannot be adequately rendered as a premise in the inference proper without appealing to material rules of inference.

Third, material inference is a philosophically more fundamental notion than formal inference, in that the former can be used to define the latter but not vice versa (*MIE*: 104-105). It is a feature of material inferences that their warrant is normally preserved upon substitution of the concepts involved in the inference. For example, if I replace every occurrence of a term (e.g., 'photon') in a materially warranted inference by another term (e.g., 'cat'), this would turn it into a materially invalid inference, as the content of the substituted terms is now different and thus licenses different inferences. However, should a material inference happen to fall under a deductively valid schema, substitution of terms does not affect its material validity, provided that no logical terms are replaced. The shift here is to see that, given the notion of material inference, we can then define the class of formal, deductively valid inferences as that class of inferences that remain *materially* valid and cannot be turned materially invalid by any substitution of non-logical vocabulary.[14]

[14] This general idea has been applied against formal schemas of inductive inference by John Norton (2003, 2010). Norton argues that all schemas that have been proposed to capture the validity of inductive inference, including Bayesianism, the error-statistical approach, and inference to the best explanation either suffer from counterexamples or else are vacuous. Norton proposes a 'material account of induction' in which all induction is local and grounded, not in the form of the

One consequence of prioritising material inference over formal inference is that it reverses more familiar conceptions of the relation between formal logic and ordinary language use. A standard assumption is that logic has its own, independent justification, which then constrains what one ought to say and believe on the basis of formal rules of inference. Brandom's account argues for a reappraisal of the role of logic and its relation to ordinary language use. As we have seen, Sellars suggested that the function of the subjunctive conditional was to express or codify an inference that we would normally just implicitly endorse in practice. Generalising this, Brandom takes the act of making inferential norms explicit in claims – paradigmatically in the form of conditionals - as part and parcel of the activity of challenging each other's commitments and entitlements in the game of giving and asking for reasons. To take a more familiar example, Brandom sees the claim, 'If Socrates is human, then he is mortal' as making explicit the propriety of the inference from 'Socrates is human' to 'Socrates is mortal'.[15] Brandom extends this idea to logic and argues that the use of logic and logical vocabulary primarily has an *expressive* role: it serves to make explicit the inferential norms and proprieties that are implicitly grasped by competent language users in the form of codified, formalised claims (*MIE*: 105-106). Logic is thus taken to acquire its justification from its ability to formalise and express in clear-cut terms the norms that are usually implicit in those practices. In that sense, the introduction of logical vocabulary to a linguistic practice provides new expressive resources for making explicit speakers' commitments, entitlements and the inferential norms endorsed by a community. It does so by articulating them as assertions which can be open to challenge and demands for reasons and justification. Thus, Brandom and Sellars argue that formal vocabulary and relations allow us to express what we must already know how to do in our practical grasp of concepts.

5.2 Material incompatibility

The second departure from the hyperinferentialist model concerns the inclusion of material incompatibility within a claim's inferential articulation; that is, which material inferences one

propositions involved, but rather in matters of fact and conceptual content that hold only in specific domains. Norton's proposal is a specific application of Sellars'/Brandom's claim that all inference should be thought of as material. For further discussion, see Brigandt (2010) and Reiss (2012).

[15] The basic move towards this view of logical vocabulary was once again anticipated by Sellars in his account of meaning as functional role (Sellars 1974). Broadly, Sellars held that 'means' claims, such as '*Rot* means "red"' do not describe word-world relations, but rather serve a normative function within ordinary language of making explicit the claim that the term on the left plays a similar functional role in the home language of the term on the right. Meaning claims thus serve as an 'inference ticket', that licenses other language-users to use the expressions in particular ways which come with certain proprieties and normative consequences.

is precluded from drawing based on commitment to certain other claims. For example, the correct applicability of 'square' precludes the correct applicability of 'triangular'; the two claims are incompatible with one another. Notice, however, that this also implies that the material inference from 'square' to '*not*-triangular' is a good one (Brandom 2007: 657). It is important for Brandom's inferentialism that incompatibility relationships between claims are included: understanding conceptual content involves a grasp not only of what ought to follow from a given claim, but also of what ought not to follow from it.

The notion of material incompatibility underwrites a particular kind of inferential relation, that of material incompatibility-entailment (*MIE*: 160). The idea here is that one claim, *p*, incompatibility-entails another claim, *q*, if everything incompatible with *q* is incompatible with *p*. To illustrate, Brandom suggest that,

(c) Wulf is a dog

incompatibility-entails

(d) Wulf is a mammal

because everything incompatible with (d) is incompatible with (c). The significance of this is as follows.

First, the reason for including the notion of material incompatibility in our understanding of conceptual content is that it extends the kind of inferential relations into which it can enter. Since we are trying to model the kind of inferences that we understand and endorse in ordinary language use, we must account for these types of relations. Doing so provides us with tools for carving out the content of our conceptual understanding by showing what our commitments preclude us from. As I suggested above however, working out what our concepts commit us to and preclude is often a 'messy retail business' (*MIE*: 601). The general point here is that the inferential consequences of our conceptual commitments typically outrun what we have a clear and transparent grasp of. What this means is that, on the whole, working out what we are committed to with using certain concepts is not something that can always be determined *a priori*, but – and particularly with respect to scientific understanding – is something that must be determined *a posteriori*. In some domains, for example in quantum mechanics, the failure to reliably determine what our conceptual commitments entail or preclude is particularly pertinent, and reflects the intuition that our understanding of quantum

phenomena is by no means secure.[16] But this need not be exclusive to the domain of quantum physics; instead, as I will argue in chapter 4, this conceptual openness is an important part of scientific research in general.

Second, incompatibility entailments put inferential meat on the normative pragmatic bones. We saw in the previous section that incompatibility was defined in terms of relations between commitments and entitlements, such that two commitments are incompatible if commitment to one precludes entitlement to the other. Transposing this into inferential semantic terms gives us an account of the way in which such normative relationships ground conceptual, semantic content.

Third, Brandom takes this kind of incompatibility relation to underwrite modal claims too, ones that can be made explicit by saying, for example: 'Necessarily, anything that is a dog is a mammal'. In effect, as we saw in the previous subsection, modal statements help to codify and make explicit the inferential norms that underpin ordinary language. Modal statements function, as Sellars would say, as inference licenses. Moreover, such statements are modally robust: if Wulf *were* a dog, then he *would* be a mammal (Wanderer 2008: 113). Thus, the fecundity of including incompatibility relations within the range of inferential articulation of conceptual content lies in the rich expressive resources it affords for bringing to light the kind of inferential proprieties already at work in the use of ordinary language.[17]

5.3 Perception, action and inference

Perhaps the most important contrast for Brandom's strong inferentialism from the narrowly restrictive hyperinferentialism is the inclusion of inferential relations between *non*-inferential circumstances of appropriate application and *non*-inferential appropriate consequences of application. Non-inferential circumstances and consequences refer to the inclusion of *perception* and *action* within our conceptual understanding of the world. For Brandom, perception and action are not specifically linguistic but nevertheless are still connected to inferentially articulated understanding.

Again, Brandom's model is Sellars, who argues that there are three kinds of moves in a inferentially articulated language game: (1) *language entry transitions*, which correspond to

[16] This point links with James Cushing's claim that quantum mechanics constitutes an 'endgame for understanding' (Cushing 1994).

[17] In an aptly-titled paper, 'Concepts as Involving Laws and Inconceivable without Them', Sellars argued that the ability to use ordinary empirical descriptive terms such as 'green', 'rigid', and 'mass' already presupposes an understanding of the kinds of counterfactual properties and relations made explicit by modal vocabulary (Sellars 1948). Below, I shall suggest that it is this feature of conceptual understanding that underpins the claim made by Grimm (2011) and others that explanatory understanding is constituted by our ability to reason counterfactually.

the ability of the agent to respond appropriately to objects in perceptual situations; (2) *intra-linguistic transitions*, in which our conceptual understanding takes place in valid material inferential moves within our conceptual repertoire; and (3) *language exit transitions*, which refer to appropriate actions in response to the particular situations in the world. The inclusion of (1) and (3), corresponding to perception and action, within our conceptual understanding is crucial to empirical conceptual content. Without connection to perceivable circumstances or practical activity, inferential relations among linguistic expressions would be empty, a "frictionless spinning in a void" in McDowell's (1994) picturesque phrase. It means that our conceptual understanding is essentially world-involving, through the discursive significance of perception and action.

To illustrate this claim, Brandom considers the difference between non-inferential reporters (sapient, rational entities, such as us) and automatic, differentially responsive detectors, such as thermostats and photocells. *Prima facie*, based on their behaviour in given circumstances, both can be classed as reliably responding to stimuli and can perhaps even be seen as classifying those stimuli as belonging to a certain kind (imagine a parrot that has been trained to say, "That's red" in the presence of red things). However, Brandom argues that the key difference is that non-inferential reporters nevertheless grasp (at least some of) the inferentially necessary consequences of applying that concept:

> […] the key element missing from the parrot and the measuring instrument – the difference between merely *responsive* classification and *conceptual* classification – is their mastery of the practices of giving and asking for *reasons*, in which their responses can play a role as *justifying* beliefs and claims. To grasp or understand a concept is […] to have practical mastery over the *inferences* it is involved in – to know, in the practical sense of being able to distinguish, what follows from the applicability of a concept, and what it follows from […] Concepts are essentially inferentially articulated. Grasping them in practice is knowing one's way around the proprieties of inference and incompatibility they are caught up in […] It is practical mastery of the inferential involvements of a response, the responder's understanding it in this sense, that makes the response an intentional state or performance – one having a content for the one whose state or performance it is, and not merely for those using it as an indicator. (*MIE*: 89)

The idea then is that even in non-inferentially reporting one's circumstances, one must understand what follows from the applicability of a concept, and what the concept itself follows from (*AR*: 48). The parrot, for instance, does not treat "That's red" as incompatible with "That's green", nor as following from "That's scarlet" and entailing "That's coloured". Similarly, the consequences of an inferential commitment also include action and non-linguistic performances (such as the connection between the claim "The traffic light is red" to the action of pressing the brakes sharply).

Brandom's inferentialist semantics thus holds that conceptual content and meaning is acquired through the use of such content in our inferential reasoning rather than in terms of a referential or representational relation to the world. The notion of material inference, which is both modally robust and incorporates incompatibility-entailments, provides a rich and expansive demarcation of what is involved in conceptual understanding. On this account, conceptual understanding can be understood not as the turning on of a Cartesian light, but as a practical mastery of the material inferential relations, norms and proprieties that underpin how we make sense of the world (*MIE*: 120).

6 Two-Dimensional Conceptual Normativity

The final element I want to consider is Brandom's account of conceptual normativity: the idea that our conceptual understanding is not simply a reliable, differential recording of the world, but is irreducibly normative. Brandom argues that conceptual normativity has an essential two-dimensional structure. The normativity of conceptual understanding means having a grasp of both the *circumstances* of appropriate application of a concept and the appropriate *consequences* of that application. As we shall see, this helps to tie together Brandom's inferentialist semantics and his normative pragmatics.

Brandom's model for this two-dimensional structure of inferential use and conceptual normativity is Dummett (1973), who proposed that the use of any linguistic expression had these two aspects: the circumstances under which it is correctly applied, uttered or used , and the appropriate consequences of its application, utterance or use. Dummett's proposal is itself a generalisation from Gentzen, who originally developed the specification of logical connectives in terms of introduction and elimination rules. Recall from above that for the connective '&', an introduction rule would state that anyone committed to *p* and committed to *q* is committed to *p* & *q*; the elimination rule would state that anyone who is committed to *p* & *q* is committed to *p* and committed to *q*. The connective is thus defined in terms of both the inferentially sufficient conditions for the use of sentences containing the connective, and in terms of the inferentially necessary consequences of the use of sentences containing the connective. Dummett's contribution was to show how this model could be extended beyond logical connectives to provide a uniform treatment of what we say and do. What corresponds to an introduction rule for propositional content is the set of *sufficient* conditions for asserting it, and what corresponds to an elimination rule is the set of *necessary* consequences of asserting (*MIE*: 116-119; *AR*: 62-66). On Dummett's model, conceptual understanding thus involves having at least some grasp of the circumstances in which it is appropriate to assert something and a grasp of what follows from the use of particular concepts.

Brandom develops this model by arguing that the inferential links – in terms of circumstances and consequences of application – between our concepts acquire content in virtue of their pragmatic significance in the game of giving and asking for reasons. Beyond grasping the inferential connections between concepts, for Brandom, the two-dimensional normativity of conceptual understanding involves a grasp of the pragmatic, normative significance that such claims and assertions have. The idea here is that asserting a sentence or a claim implicitly involves 'undertaking a commitment to the correctness of the material inference from its circumstances to its consequences of application' (*AR*: 63). That is, our expressions acquire conceptual content not only in virtue of their inferential relations, but also in virtue of their corresponding moves in the game of giving and asking for reasons, of the shifts between our commitments and entitlements.

In Section Three, we saw that Brandom cannot defend his normative inferentialism by appealing to the notion of truth. Since he is committed to understanding truth in inferentialist and pragmatic terms, he cannot understand materially good inferences in terms of the preservation of truth in the move from premises to conclusions. Instead, Brandom understands inference in broadly pragmatic and non-semantic terms: valid or good inferences are understood not in terms of the preservation of truth, but in terms of *preservation of normative status*. On this construal, a good inference is one which preserves one's entitlement to a commitment. For example, avowing a commitment to *q* in response to a challenge to *p* is a 'good' or 'valid' move if and only if entitlement to *p* preserves entitlement to *q*. Hence, the notion of materially good inference depends for Brandom on the normative attitudes of scorekeepers situated in social, linguistic practices. This is why Brandom wants to understand inference not in terms of dispositions to infer on the part of subjects, but in terms of norms of *correct* and appropriate inferences: the ones people *ought* to make, rather than the ones they are actually disposed to make (Brandom 2007: 656).

Brandom argues that it is a failure to distinguish between the circumstances and consequences of application of concepts that leads to 'one-sided' theories of meaning. The mistake of verificationists, assertibilists and reliabilists is that they only treat the *first* aspect – circumstances of application – as exhausting content. For the assertibilist and verificationist, understanding conceptual content is taken to consist in practically mastering the circumstances under which one becomes entitled or committed to endorse a claim, quite apart from any grasp of what one becomes committed to by such endorsement. But this cannot be right:

> […] claims can have the same circumstances of application and different consequences of application, as for instance, "I foresee that I will write a book about Hegel" and "I will write a book a about Hegel" do. We can at least regiment a use of 'foresee' that makes the former sentence have just the same assertibility conditions as the latter. But substituting one for the other turns the very safe conditional "If I write a book about Hegel, then I will write a book

> about Hegel," into the risky "If I *foresee* that I will write a book about Hegel, then I will write a book about Hegel." The possibility that I might be hit by a bus does not affect the assessment of the inference codified by the first conditional, but is quite relevant to the assessment of the second inference. (*AR*: 64)

Similarly, against the reliabilist, although Brandom endorses the shift towards an externalist form of justification, in which part of one's entitlement to a claim consists in one's ability to respond reliably to particular circumstances (*AR*: 97-102), he argues that reliability alone is not enough to confer content on our claims. What is further required is a grasp of the inferential consequences of one's claims and their normative significance in discursive, social practice.[18]

More generally, however, the significance of this two-dimensional structure for conceptual understanding and normativity is that it mirrors the fact that what we say and do – our thoughts, actions, claims, and performances – are normatively accountable in two distinct, but interconnected ways. These two dimensions correspond to their meaningfulness on the one hand, and their justification or truth on the other. It thus introduces a possible gap between how we *take* the world to be, and how the world actually *is*. The significance of this gap is that it allows for the possibility of *error*.[19] As such, genuine conceptual understanding involves both the ability to determine, on the one hand, whether what has been said is meaningful, intelligible or significant, and then on the other hand, whether what has been said is justified, warranted, true, or more broadly, should be the kind of claim that the social practice or community should endorse. It is because of these two-dimensions to the normative accountability of conceptual understanding that revisions and shifts in our understanding of the world become possible, because only with a prior grasp of what a particular claim says about the world can we then be in a position to reflect upon and evaluate whether it is true or justified.[20] It is this dual dimensionality to our conceptual understanding that enables our

[18] In Brandom's words: 'You do not convey to me the content of the concept *gleeb* by supplying me with an infallible gleebness tester which lights up when and only when exposed to gleeb things. I would in that case know what things were gleeb without knowing what I was saying about them when I called them that, what I had found out about them or committed myself to' (*AR*: 65).

[19] For further discussion of this point, see Haugeland (1998: 314) and Rouse (2015a: 253-254).

[20] For various neo-pragmatists, such as Brandom, Davidson, or Haugeland, it is this two-dimensionality to the normativity of conceptual understanding that distinguishes the rationality of human organisms, from the goal-directed behaviour of non-human organisms. The point here is that, although non-human organisms can exhibit a wide range of attuned and flexible responsiveness that is highly successful and appropriate in response to its particular environmental milieu, the correctness or appropriateness of their behaviour is not something that can become an issue for them. What is exhibited is a one-dimensional normativity in the sense of whether or not they are successful at achieving their particular goals, but lacks the two-dimensionality to determine whether the particular behaviour they undertake is itself appropriate to the achievement of those goals. What distinguishes us is that not only does our conceptual understanding allow us

inferential commitments and entitlements to be open to challenge, refinement and clarification. This in turn underpins what is at stake in conceptual change. Thus, Dummett writes, conceptual change can be,

> […] motivated by the desire to attain or preserve a harmony between the two aspects of an expression's meaning. A simple case would be that of a pejorative term, e.g. 'Boche'. The conditions for applying the term to someone is that he is of German nationality; the consequences of its application are that he is barbarous and more prone to cruelty than other Europeans. We should envisage the connections in both directions as sufficiently tight as to be involved in the very meaning of the word: neither could be severed without altering its meaning. Someone who rejects the word does so because he does not want to permit a transition from the grounds for applying the term to the consequences of doing so. (Dummett 1973: 454)

This passage is significant for the following reasons. First, it highlights how concepts connect and how they can be criticised on the basis of one's prior commitments. One can refuse to employ the concept (e.g., 'Boche') on the grounds that it embodies an inference one does not endorse (*AR*: 70). Second, it points to the important function of critical thinking, which can be understood as an attempt to make explicit (often through formal logical vocabulary) potentially controversial material inferential commitments that should be scrutinised. Third, the basis of working out our conceptual commitments and the circumstances and consequences of their application can provide the basis for the introduction of novel content and inferential norms. This would be contrary to Dummett's invocation of the need to preserve harmony between inferential relations, but is important when accounting for conceptual progress in science. For example,

> […] the concept of temperature was introduced with certain criteria or circumstances of appropriate application, and certain consequences of application. As new ways of measuring temperature are introduced, and new theoretical and practical consequences of temperature measurements adopted, the complex inferential commitment that determines the significance of using the concept of temperature evolves. (*AR*: 71; *MIE*: 127)[21]

This highlights the point that the material inferential commitments endorsed by agents and the communities in which they are embedded should not be seen to be rigid or fixed, but open to revision and articulation through the process of ongoing inquiry. Furthermore, determining

---

to adjudicate or judge the appropriateness of a particular claim or performance in a situation or practice, but also the broader significance of that claim or performance.

[21] Hasok Chang's (2004) study of the concept of temperature can be read in this light, as detailing the competing shifts within the scientific communities of the 17th and 18th centuries concerning what the appropriate circumstances and consequences of application for 'temperature' were. Importantly, Chang's project shows that such inferential articulation is not simply an intralinguistic affair, but one that incorporates the material, concrete dimension of experimental cultures.

the normativity of our conceptual understanding - working out which particular claims about the world to endorse and the inferential rules that they institute - is not something that can be resolved independently of a particular practice, but can only be sorted out from within. This means that conceptual normativity has an inherently *perspectival* nature. There is, as it were, no single perspective which is in principle inherently privileged or dominant or has an automatic claim to entitlement. Thus, Brandom suggests that:

> Each perspective is at most *locally* privileged in that it incorporates a structural distinction between objectively correct applications of concepts and applications that are merely subjectively taken to be correct. But none of these perspectives is privileged *in advance* over any other […] Sorting out who *should* be counted as correct, whose claims and applications of concepts should be treated as authoritative, is a messy retail business […] [T]here is no bird's-eye view above the fray of competing claims from which those that deserve to prevail can be identified. (Brandom 1994: 600-601; italics added)

Questioning our inferential commitments and attempting to clarify their consequences highlights their fundamental normative significance: in this case, our understanding concerns not simply which inferences one is capable of making or endorsing, but also, more substantively, an appreciation of the issue of which inferences one *ought* to endorse. In the following chapter, I will develop an account of norms that is able to accommodate the dynamic nature of conceptual progress, but one which, more significantly, places the ongoing issue of what inferential commitments one ought to endorse at the heart of conceptual progress, not only in scientific research and inquiry, but in our discursive practices more generally.

7 Concluding Remarks

Brandom's inferentialist semantics, combined with his normative pragmatics, offers a rich and complex account of conceptual understanding. At the heart of this account is the contention that inferential ability by itself is not enough to capture what is at stake in conceptual understanding. Instead, the significance of our inferential abilities and what we say and do is tied to the normative structure of our social, discursive practices. Understanding as inferential ability is not simply about what we can do with particular conceptual content, but more substantively, is underpinned by our responsiveness to the norms embedded within our practices. In the next chapter, I will focus on the nature of these norms and how they structure our practices. To end this chapter, I will consider the implications of Brandom's account of conceptual understanding for the current debates on understanding.

First, Brandom's view endorses the claim that understanding is constituted by our inferential ability. Like many authors in the debate, who have offered different accounts of

what is involved in 'grasping', Brandom would agree that merely believing or knowing something does not qualify as understanding. Instead, one must be able to do something with that content; one must have a kind of practical know-how in relation to that knowledge. This concerns the inferential articulation of our concepts. Not only must one have a grasp of how and where concepts apply, one must also have a sense of what follows from those concepts as well as the consequences of their application. For Brandom, our inferential ability is an essential feature of what it means to be a discursive, rational animal. Conceptual understanding is constituted by the ability to use concepts in our reasoning, to apply them in appropriate circumstances and to have a sense of what follows from that use. Significantly, however, Brandom departs from the current literature by holding that this inferential ability is not a narrowly psychological or cognitive capacity on the part of individual knowers, but rather constitutive of our 'being-in-the-world' as norm-responsive agents, situated in intersubjective spaces of meaning.

Second, characterising understanding in terms of our inferential abilities and the practical mastery of the inferential consequences of our concepts leads to the idea that understanding comes in degrees. For Brandom, we can have greater or lesser understanding depending upon our training, education and expertise and that different levels of expertise reflect different levels of conceptual understanding. For example: 'metallurgists understand the concept *tellurium* better than most of us do, for training has made them master the inferential intricacies of its employment in a way that we can only crudely approximate' (*MIE*: 120).

Third, Brandom's incorporation of material inference as the central aspect of our inferential ability helps to clarify and articulate the *kinds of inferences* that one must be capable of grasping in order to qualify as understanding. First, the appeal to material inference suggests that the inferences constitutive of understanding are not primarily formal in character. Second, the notion of incompatibility suggests that one understands also involves a sense of which inferences one is precluded from drawing. Narrowly, we might think of this in terms of logical incompatibility, but more broadly, the kinds of inferential incompatibilities involved will also be accountable to the inferential norms endorsed by the community.

This in turn leads to the fourth consequence: the modal significance of our inferential abilities. As we have seen, various authors (Grimm 2006; Hills 2015; Woodward 2003) endorse the idea that the abilities constitutive of 'grasping' incorporate our ability to reason counterfactually. Others, however, have challenged this idea, claiming that it is too demanding to be constitutive of understanding (e.g., Newman 2013). Brandom's and Sellars' arguments for the priority of material inference offer a response to this challenge and clarify the scope of the original point. Their arguments showed that ability to use even our ordinary, empirical concepts – concepts such as 'red', 'rigid', 'acid' – already presuppose an understanding of the kinds of properties and relations that can be made explicit in the form of subjunctive

conditionals (Brandom 2008: 97). To understand a concept, for Brandom (and Sellars) means having an appreciation of the counterfactual possibilities that underpin those concepts. For example:

> One grasps the claim "the lioness is hungry" only insofar as one takes it to have various consequences (which *would* be true if it *were* true) and to rule out some others (which *would not* be true if it *were* true) […] The claim is not that one could not fail to assess some, or even all, of *these particular* counterfactuals correctly and still count as grasping the claim that it is their premise, but that one could not so qualify if one made *no* such distinctions. (ibid: 105)

The upshot of Brandom's point here is that the ability to consider counterfactual cases – what *would* happen if *x* were the case – is not something that one acquires on top of some presumed minimal level of understanding. Instead, it is essential that one has a grasp of the kinds of possibilities that one's inferential commitments are open to in order to count as having conceptual understanding as such.[22]

Finally, the crucial contribution I take from Brandom's account is that it shows how conceptual understanding is irreducibly normative. This is not only in the sense that our inferential abilities are governed by rules of material inference, which underpin the move from the circumstances of appropriate application of our concepts to their appropriate consequences. More generally, it means that our understanding has meaning and significance in light of its relation to the normative pragmatics of what we say and do in our norm-governed social practices. From this perspective, conceptual understanding involves more than simply representing the world in the right kind of way; more crucially, it is dependent on the way in which our inferential abilities and claims about the world are accountable to a shared way of life and the norms of intelligibility endorsed by particular communities.

[22] As I read it, the importance of alethic modalities in Brandom's account of conceptual understanding represents an important elaboration of Heidegger's more general claim that understanding the being of entities always involves a 'projection onto their possibilities'. Heidegger's account of ontological understanding, in which we understand things in terms of what is possible or not possible for them, would thus suggest an anticipation of Brandom's more general point. Arguing for this claim, however, will have to the subject of future work.

Chapter Three

Rules, Norms, and Practices

1 Introduction

My discussion of Heidegger and Brandom in the previous chapters highlighted the connection between understanding and normativity. For both thinkers, our understanding is constituted and sustained by the normative structure of the concrete, social practices in which we, as subjects of understanding, are embedded. This chapter addresses the concept of normativity itself and its role in constituting and shaping human practices. In the thesis introduction, I defined normativity in a broad sense as a covering term for any phenomenon that serves as a standard for success or failure of any kind. A norm by this definition can be anything that serves as a measure for what we think, say, or do. The aim of this chapter is to expand upon and clarify this definition.

In this chapter I begin by critically examining two prominent and intuitive philosophical accounts of norms and normativity: *regulism*, which treats norms as explicitly formulated, prescriptive rules, and *regularism*, which treats norms as regularities in behaviour.[1] Both conceptions are shown to be inadequate as accounts of normativity. Instead, I introduce and defend a conception of norms as interlocking sets of normative attitudes. The key aspect of this alternative conception is that participants in any social practice are mutually accountable to one another. I then draw upon Joseph Rouse's *temporal normativism* to introduce the idea that human practices are *temporally extended* patterns of behaviour, whose normativity emerges from something being continuously at issue and at stake in those practices. From this, I argue that the conceptual normativity that constitutes and shapes our understanding is grounded in situated, discursive and temporally extended practices.

2 Regulism: Norms as Rules

Wittgenstein's discussion of rule-following in the *Philosophical Investigations* offers one starting point for our understanding of the concept of normativity. The primary focus of

[1] Regulism and regularism are terms coined by Brandom (*MIE*: 18-26).

Wittgenstein's discussion is the normativity of linguistic meaning, but we can understand Wittgenstein's arguments to be directed against certain assumptions about the normative character of human thought, action and performances in general. In the *Philosophical Investigations*, heexplores the idea that what could give a word its meaning is a rule for its use, and that to be a competent speaker is to use words in accordance with these rules. Kripke's interpretation of Wittgenstein's rule-following discussion concludes that an adequate account of meaning and intentional content must not simply describe actual behaviour or dispositions to behave. Rather, 'the relation of meaning and intention to future action is *normative* and not *descriptive*' (Kripke 1981: 37). The claim here is that an account of the content of what we say or do must say something about how people *ought* to behave if they think this or mean that. For example, when one employs the concept of addition in order to answer the question of what the sum of 2 and 19 is, it is reasonable to assume that one *ought* to come to the conclusion that it is 21. Similarly, when one judges that something is red, one *should not* go on to deny that it is coloured. Despite this, the strands of discussion that make up Wittgenstein's remarks on rule-following pose fundamental concerns, as we shall see, for any conception of normativity that understands it primarily in terms of rules.

In certain respects, an appeal to rules seems to be a perfectly natural one, because rules are just the sort of things naturally taken to determine what one ought to do, and it seems plausible that we often learn such rules in learning the meaning of a term like 'plus'. It seems natural to suppose that if we are to follow a rule and act with an understanding of it, the rule should be formulated in some way in order to express what should or should not be done. Without such a formulation, how would we know whether a word has been used correctly? It thus appears to be plausible that acting with an understanding of rules is an essential requirement for meaning and communication to be possible. Nevertheless, Wittgenstein's discussion suggests that the matter is not so simple.

Wittgenstein's remarks on rule-following are primarily directed against a particular conception of the nature of rules and norms in general, which Brandom terms *regulism* (*MIE*: 18-26). Regulism identifies any norm on the model of prescriptive rules that (explicitly) specify the actions that accord with them. On this model, a regulist conception of human practices would identify practices by a set of rules which govern their practitioners. Despite the intuitive nature of this conception of norms and practices, Wittgenstein wishes to criticise the assumption that the normative force of the use of such concepts be identified with verbally or conceptually formulated rules, or more specifically, that a rule should be conceived as some kind of mental content that one has before one's mind. This would at the very least threaten the idea that practices can be thought as demarcated by sets of rules. The central question that Wittgenstein identifies is: how does the presence of such a rule *commit* one to a given behaviour?

The crux of the problem is that rules are not self-interpreting entities; they do not, by themselves, determine what a correct or incorrect application of the rule is. How we apply a given rule to the present situation will depend on how one *interprets* that rule. However, for any given rule, there are any number of interpretations one could apply to it. This would seem to require the existence of a further rule that would determine the correctness of the initial interpretation and pin down which interpretation one should follow. But here we find ourselves on the slippery slope of a vicious regress because an interpretation has exactly the same normative force as the original concept insofar as it serves as a rule prescribing certain behaviour. The result is then the need for an interpretation of the interpretation, and so on.[2] The resulting point that the rule-following considerations highlight is that any conception of the normative that is modelled on the existence of some presented rule makes it fundamentally mysterious how such an entity could ever prescribe how one ought to behave. From these considerations, Wittgenstein drew a complex conclusion:

> This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it also can be made out to conflict with it. And so there would be neither accord nor conflict here. It can be seen that there is a misunderstanding here from the mere fact that in the course of our argument we give one interpretation after another; as if each one contented us for at least a moment, until we thought of yet another standing behind it. What this shows is that there is a way of grasping a rule which is not an interpretation, but which is exhibited in what we call 'obeying a rule' and 'going against it' in actual cases. (*PI*: §201)

The upshot of Wittgenstein's criticism is that there must be a level or dimension of human thought and action expressed in what we do that is more fundamental than any explicit interpretation of that understanding, 'a way of grasping a rule which is not an interpretation'. As Brandom concludes, 'there is a need for a pragmatist conception of norms – a notion *of* primitive correctnesses of performance *implicit* in practice that precede and are presupposed by their explicit formulation in rules and principles' (*MIE*: 21; emphasis mine).[3] If, as the regulist holds, to act according to norms is to follow a rule, and rule-following can be done correctly or incorrectly, then a vicious regress of rules would render action according to norms

[2] A linguistic version of this argument finds expression in Sellars', 'Some Reflections on Language Games'. There Sellars argues that if we construe language as a system of expressions for which there are rules of use, this would presuppose an ability to use a metalanguage in which the rules for the original language are formulated. But similarly, the rules for using the metalanguage would then presuppose a meta-metalanguage, and so on, leading to the regress (Sellars 1954: 204).

[3] McDowell (2009) disputes Brandom's interpretation here, contesting the claim that any such argument against a regulist conception of norms can be found in Wittgenstein. However, I will leave the issue of interpretive authenticity to one side.

impossible. Because regulism generates this regress, it must be rejected as a possible account of practices and the norms governing them, as on this conception both would be rendered incoherent. The challenge, then, is to characterise a way of obeying rules without interpreting them and which is exhibited 'in actual cases'.

Here, Wittgenstein famously goes on to say, 'That's why "following a rule" is a practice. And to *think* one is following a rule is not to follow a rule' (*PI*: §202).[4] In this statement, the concept of a 'practice' is invoked because it is taken to identify the locus of our background understanding and competence which is exhibited in rule-following behaviour, but without requiring an explicit interpretation of those rules. This would seem to be an essential condition if we are to preserve the normative dimension of our thought and action that the regress argument renders problematic. As Brandom notes, 'Absent such a practical way of grasping norms, no sense can be made of the distinction between correct and incorrect performance – of the difference between acting according to a norm and acting against it' (*MIE*: 21). Wittgenstein's criticisms of regulist conceptions of normativity are thus propaedeutic to a more appropriate, alternative conception of practices. The challenge then is to articulate a notion of rule-following and acting from an understanding of a rule in practice in a way that would halt the regress of rules and interpretations.

3 Regularism: Norms as Regularities

One way of avoiding the regress would be to question some of the assumptions that seem to be at work in the picture of rule-following that generates it in the first place. In Wittgenstein's initial presentation of the problem, following a rule seemed to require the possibility of having an interpretation for the application of that rule. But is it the case that following a rule *R* always requires following another rule R^1 for *R*'s proper application? Challenging this claim would be one way to resist the regress. As McDowell puts it, we should reject the claim that 'grasping a rule is always an interpretation' (McDowell 1984: 338-9).

The basis for this challenge is that there seem to be plenty of occasions in everyday life where we can and seem to correctly follow a rule without recourse to an interpretation of that rule or the need for a further rule. When I follow a road sign indicating the direction to my destination, I do not (normally) have to interpret the road sign in order to determine what it would mean; I just know how to follow it. Normally, no question arises for us about what it means so that we would not need to consult a rule for what to do in such circumstances. According to Brandom, the picture of rule-following that seems to be at work in the regress is

[4] Similarly: 'To follow a rule, to make a report, to give an order, to play a game of chess, are *customs* (usages, institutions).' (*PI*: §199).

thus overly 'intellectualist'. It seems to demand the implausible scenario that understanding a rule requires us to 'think' and 'interpret' any time we manage to follow a rule or abide by norm (*MIE*: 20). But this intellectualist assumption is undermined by the kind of 'skilful coping' exhibited by most ordinary behaviour.

In the light of this line of thought, a common and natural response to the regress problem is to move from an intellectualist and mentalistic account of rule-following to a behaviouristic one (Lance and O'Leary-Hawthorne 1997: 182). Such a suggestion is introduced by Sellars:

> Now, at first sight there is a simple and straightforward way of preserving the essential claim of the thesis while freeing it from the refutation. It consists in substituting the phrase 'learning to *conform to* the rules…' for 'learning to obey the rules…' where 'conforming to a rule enjoining the doing of A in circumstances C' is to be equated simply with 'doing A when the circumstances are C' […] A person who has the habit of doing A in C would then be conforming to the above rule even though the idea that he was to do A in C had never occurred to him, and even though he had no language for referring to either A or C. (Sellars 1954: 204)

Thus, the response that Sellars articulates is to suggest that the correct application of a term is determined not by a mentally grasped rule, but by conformity with a behavioural habit acquired in learning to use the term. The basis for such a response seems to be countenanced in Wittgenstein's remark that, 'If I have exhausted the justifications [for following a rule in the way I do], I have reached bedrock, and my spade is turned. Then I am inclined to say, "This is simply what I do"' (*PI*: §217). Moreover, Wittgenstein goes on to characterise this habitual action as unreflective: 'When I follow the rule, I do not choose. I follow the rule *blindly*' (*PI*: §219). A similar suggestion seems to be at work in *Being and Time*, where Heidegger describes our everyday understanding and interpretation as structured by what 'one' or 'the they' [*das Man*] does, which likewise evokes some level of pre-reflective conformity in behaviour.

Moreover, this behaviourist approach would appear to provide a way of accounting for the *appearance* of normativity in our thought and action by explaining such thought and action in terms of non-normative regularities in behaviour: one follows a rule by acquiring a habit to do what the rule requires in the right circumstances. Concluding that the sum of 2 and 19 is 21 is, on this conception, a habit or behavioural regularity acquired through training and education to respond to these terms in a specific way. In this respect, a non-normative domain of behavioural regularities can be seen to sustain our normative practices.

Brandom labels this proposal *regularism* (*MIE*: 27). According to regularism, the regress of rules comes to an end in a behavioural regularity *exhibited* by what practitioners do, rather

than in a rule *followed* by them.[5] Practices are thus conceived in terms of shared dispositions and regularities in behaviour, and to invoke normativity here is to speak of the habits and dispositions to behave in some ways rather than others. The correctness of any given performance depends on the extent to which it is in accord with a behavioural regularity. Thus, whereas *regulism* models practices and the norms that govern them on explicit, prescriptive rules and the interpretation of those rules, *regularism* models them on commonalities of performance and regularities in behaviour.

Regularism is an attractive idea. First, it would seem to sidestep the crucial assumption that generates Wittgenstein's regress, namely that following a rule always requires following another rule for its application arrived at through an interpretation. Second, it accords with the phenomenology of everyday behaviour, one which does not seem to require the explicit interpretation of rules in order to correctly follow them, and to thus do away with the intellectualist character of the original regress problem. Third, and more ambitiously, the appeal to behavioural regularities in order to explain practices and their norms could be seen to provide a way of explaining them in causal, naturalistic and fundamentally non-normative, non-intentional terms. By explaining the normativity of practices in non-normative terms, as regularities and habits of behaviour, such a strategy would have the advantage of providing a straight-forward, non-circular account of normativity.[6] This naturalistic explanation of rule-following behaviour and normativity could purportedly claim to provide a non-mysterious, natural grounding for normativity Despite this, the regularist conception of practices and norms suffers from several intractable problems that render it unacceptable.

First, we can begin by noting that exhibiting a regularity does not count as *following* a rule in the sense that that rule has any normative force for me. Following a rule is more than just a regularity in what we do. As Jaroslav Peregrin observes, we 'do not want to admit that there is no difference of principle between a car driver following the rules of traffic and a stone following the law of gravity' (Peregrin 2014: 70). Thinking of norms as regularities in behaviour does not seem to capture this extra dimension of rule-following.

A second problem is that our behaviour, actions and performances, verbal or otherwise, do not exhibit any *single* regularity or pattern; what we say and do instead exhibits a *variety* of regularities. One's past performances are consistent with an *infinite* number of ways of continuing that performance. Therefore, if there is no single pattern that one's performances

[5] By the above remarks concerning the philosophical background for this view in Wittgenstein and Heidegger, I do not mean to say that I take them to have endorsed a regularist conception of norms and practices. This is one way in which they can be read, but one that I think would be mistaken.

[6] That normativity should be explicated in non-normative terms on pain of a vicious regress is a common objection in the normativity of meaning debates. For example, see Hattiangadi (2006: 235).

exhibit, then those performances leave it indeterminate which rule one is allegedly following. If that is the case, then there is no way of determining what is and is not an error (Maher 2012: 49). The objection is that regularities in behaviour will fail to sort performances into correct ones (those that are in accord with the regularity) and incorrect ones (irregular). Grounding the norms governing practices in regularities provides no principled ground for selecting one regularity over others. One symptom of this difficulty – and a crucial objection to regularism – is the problem of gerrymandering (*MIE*: 28). This is the problem that, for any finite set of regularities in performance, it is possible to construct an arbitrarily large number of rules for which those performances would be an instantiation. As a result, when presented with any behaviour that deviates from a previously articulated rule, it is always possible to generate some other rule, with which that behaviour would be consistent.

Third, regularism is subject to a further problem identified by Kripke. The defender of regularism, to circumvent the gerrymandering problem, might appeal to *dispositions* to act, instead of simply regularities with one's past performances in order to placate this worry. The shift would be from arbitrarily locating the correctness of any performance in one kind of pattern of behaviour to locating it in a disposition acquired in learning the use of a given term. This would seem to address the gerrymandering problem since dispositions are dispositions to specific behaviours and would thus seem to provide a way of specifying which regularity is the *right* one. According to this proposal, one follows the rule when one does what one is disposed to do and one violates the rule when one does not do what one is disposed to do.

The problem with this proposal is that it provides only the semblance of normativity rather than genuine normativity. Dispositions do not settle what one should or should not do (Maher 2012: 50). Kripke famously argued that while a disposition may determine certain behaviour, the determination will be causal rather than normative. The problem is that if we identify what one should or ought to do with some dispositional, causal process, then anything one does will necessarily come out as correct, which is to say that, ultimately, the distinction between correct and incorrect would be vacuous. The issue is that an appeal to dispositions provides only a descriptive account of what people actually do or are disposed to do: 'to be disposed to use a word in a certain way implies at most that one *will*, not that one *should*' (Boghossian 1989: 513). It does not have the resources to tell us what we ought to do, and thus be a genuinely normative account of meaning and rule-following behaviour. As such, neither appealing to regularities with past behaviour, nor to dispositions to future behaviour will help the normative theorist.

Let me summarise the discussion up to this point. I began by considering Wittgenstein's rule-following discussion, which criticises *regulist* conceptions of norms, according to which norms are primarily prescriptive rules and practices are accordingly identified by sets of rules. This led to a vicious regress since any application of a rule seemed to require an interpretation

to determine the correctness of that application, but this interpretation itself required an interpretation, and so on. A plausible response to this regress was to challenge the intellectualist assumption which held that any application of a rule requires an interpretation of that rule. Instead, we observed that much of our everyday rule-following behaviour does not depend on having an interpretation of a rule; instead, we seem to just know how to follow a rule. Thus, rule-following could be taken simply to consist in regularities in our behaviour, rather than in terms of explicitly following content present to our minds. As such, the *regularist* proposal avoids the regress by rejecting the intellectualist assumption. However, we saw that it did so at the cost of the possibility of genuinely following rules at all. This fails to preserve a normative distinction between doing what one ought to do from what one just normally does, and of accounting for the possibility of error. Thus, while Wittgenstein's criticisms of regulism remind us not to confuse norms implicit in our practices with explicit formulations of them as rules, the anti-intellectualist, regularist response seems to swing too far in the opposite direction and consequently loses sight of the distinction between what *is* done and what *ought* to be done.

4 Normative Attitudes

What is the difference between a genuine rule or norm, on the one hand, and a complex regularity of behaviour on the other? What is the difference between my following the rules of traffic and a stone obeying the laws of gravity? One aspect that seems to be missing from the regularist conception of norms can be illustrated by a distinction. From a certain perspective, following a rule can be looked at as a complex regularity. But a rule that one ought to do X is never just a matter of doing X regularly. As Peregrin notes, following a rule to do X also involves other behavioural patterns and attitudes to those who do not do X, perhaps in the form of sanctions, disapproval, or rejection. In that respect, following a rule or norm involves more than just a regularity in behaviour; a rule is something that can be '*accepted*, *endorsed*, or *participated in*' (Peregrin 2014: 70). This suggests that to fully capture what it means for something to be normative, we need to account for how a rule or norm is made sense of by someone within a practice. The relevant distinction here is between *external* conceptions of norms and *internal* conceptions of norms; that is, norms viewed from an outsider perspective, and norms that are in force for members of a practice or community.

An account that points towards this distinction is suggested by H.L.A Hart in *The Concept of Law*. Hart proposes that the difference between norms and mere habits in behaviour is that norms involve a 'reflective critical attitude' to certain behaviour in light of normative principles:

> Chess players do not merely have similar habits of moving the Queen in the same way which an external observer, who knew nothing about their attitude to the moves which they make, could record. In addition, they have a reflective critical attitude to this pattern of behaviour: they regard it as a standard for all who play the game. Each […] 'has views' about the propriety of all moving the Queen in that way. These views are manifest in the criticism of others and demands for conformity made upon others when deviation is actual or threatened, and in the acknowledgement of the legitimacy of such criticism and demands when received from others. For the expression of such criticisms, demands, and acknowledgements a wide range of 'normative' language is used. 'I (You) ought not to have moved the Queen like that', 'I (You) must do that', 'That is right', 'That is wrong'. (Hart 1961: 55-56)

Hart's suggestion then is that norms are *clusters of normative attitudes*: patterns of behaviour in which we hold something to be correct or incorrect. We can thus say that although the regulist and regularist conceptions of norms offered *prima facie* plausible *external* conceptions of norms and rule-following behaviour, they ultimately failed to address *what it is like* to follow a rule, for a rule to be in force *for us* as participants within a practice. For an outsider, normative attitudes might well display themselves as complex regularities in what we say and do, but for an insider, what we say and do is always more than just a regularity in our behaviour. The suggestion then is that norms are neither explicitly formulated rules (on pain of regress), nor mere behavioural regularities. Rather, I submit that norms should be conceived of as clusters of normative attitudes. In what follows, therefore, when I refer to rules or norms, they should be thought of along these lines.

How should we understand norms *qua* normative attitudes? One implication of this view is developed by Peregrin. Peregrin addresses an objection faced by accounts that treat what we say and do as essentially rule-governed behaviour, which argues that rules are incompatible with the obvious spontaneity, freedom and creativity that distinguishes human lives. At issue here is the idea that following a rule or being governed by a norm means that the rule essentially dictates what to do, and hence to be governed by that rule is to be unable to act freely or spontaneously. Peregrin argues, however, that this mischaracterises what is going on in rule-following behaviour. He argues that the role of (implicit) rules and norms in human practices is not *prescriptive* in the narrow sense that they tell us what to do or guide us in our behaviour, but rather *restrictive*: rules and norms tell us what *not* to do, what is *prohibited* (Peregrin 2014: 72). The idea is that under the former conception, we are not left with much space for deliberate choice, whereas with the latter we are. From this perspective, rules and norms need not be explicitly formulated, but their normative force is still in place through the responsiveness of participants within a particular practice to their own and each other's' behaviour. This is not to deny that we can and often do have explicit prescriptions for behaviour in place. The suggestion is that these are special cases, and the injunction to say and

do something is derivative from a more general sense in which rules restrict what it is possible to do without necessarily dictating what we do.

What is the significance of this distinction? For one, it corresponds with the material inferential norms that were discussed in the previous chapter with regards to Brandom's inferentialism. As we have seen, Brandom's account takes linguistic, discursive practice to be an interactive field of performances, mediated by each participant's implicit tracking of the commitments and entitlements accrued by the various participants. Each performance affects the significance of what others mean by changing the 'score' of commitments and entitlements accrued to each participant – what Brandom calls 'deontic scorekeeping'. Here, the central linguistic performance for Brandom's account is that of asserting. For Brandom, asserting is fundamentally normative and interpersonal: to assert something is simultaneously to undertake a justificatory responsibility to defend what one has asserted, and to issue a re-assertion license to others of what one has asserted. Within this linguistic 'game', players may challenge, rely on, and defer to the assertions of one another. If an assertion is challenged, the player who made it is obliged to defend it with another assertion, to point to some other player who can defend it, or to withdraw it (Maher 2012: 67). For example, if you assert that Fido is a dog, you are not obliged to assert that Fido is an animal, but you *are* obliged not to deny it. If you do deny it, you then become a 'legitimate target of criticism' (ibid: 72). In this respect, it is the restrictive nature of rules and norms, as opposed to a narrower prescriptive nature, that more accurately underpins the meaningfulness and intelligibility of what we say and do. Rules '*constrain* us: if we want the sounds we emit to count as meaningful utterances, these rules dictate to us to avoid certain ways of using them' (ibid: 73). In that respect, the establishment and operation of a systems of norms, as clusters of normative attitudes, can be thought of as creating and delimiting a 'space of meaning' in which what we say and do can be intelligible and meaningful at all. As subjects of understanding, we inhabit such a space of meaning, which is a thoroughly normative space insofar as the rules and norms that govern it are in force or binding for us. This in turn would reflect Heidegger's more general phenomenological characterisation of the subject of understanding as a subject that is situated in a world. For Heidegger, the world is not just a collection of entities, but a horizon of meaning.

Norms and rules thus exist and are sustained through normative attitudes. These normative attitudes represent a tacit commitment to engage with others and the world in a specific way, one that is directed towards how the world ought to be rather than how it simply is. Importantly, these normative attitudes may be expressed in particular kinds of speech acts, which Peregrin calls *normatives* (ibid: 77). Normatives are claims to the effect that something is correct or incorrect, or that something ought to be done, for example, 'Doing this would be wrong'. In this respect, they are different from assertions or reports: they do not simply

describe what happens to be the case, but instead state that something *ought to be* the case. More specifically, normative claims are primarily understood on the internal, insider conception of norms: they are a way in which the utterer can be seen as endorsing a particular norm or rule. Normatives should not be understood in an epistemic, moral or political sense. Instead, what is at issue here is the very meaning and intelligibility of what we say and our actions and performances in the world. The idea is one originally proposed by Sellars with regards to semantic vocabulary. Broadly speaking, Sellars held that 'means' claims, such as '*Rot* means "red"', do not describe word-world relations, but rather serve a normative function within ordinary language of making explicit the claim that the term on the left plays a similar functional role in the home language of the term on the right. Meaning claims thus serve as 'inference tickets', that license other language-users to use the expressions in particular ways (Sellars 1974).

For Peregrin (and Sellars) normatives do not occur in isolation. Instead, their normative force derives from their being located within a social practice. As such, issuing a normative claim is possible only in the context of some practices that are already rule-governed. Lance and O'Leary Hawthorne elaborate on the idea of normatives in the following way

> Normatives are in many ways just like ordinary declaratives. They take their place in the game of giving and asking for reasons, serving as premises and conclusions in reasoning. In some respects then, normatives have criteria and consequences of application that are like declaratives; they follow from certain claims and certain claims follow from them […] But in another crucial respect their consequences of application are like imperatives. Like a declarative, one of the consequences of application of a normative is entitlement to certain other claims. But unlike declaratives, one of the direct, and widely stable, consequences of application of a normative is the appropriateness of some act: to commit oneself to a normative is *ipso facto* to commit oneself to the propriety of some act. (Lance and O'Leary-Hawthorne 1997: 202-203)

We can think of normatives as injunctions made on the part of individuals within a practice or community to other members of that practice or community; those for whom the normative would have the appropriate force. But normatives are not always just the products of individual claims about how things ought to be; instead, Peregrin argues that genuine normatives reflect what Martin Kusch calls 'communal institution-creating performatives':

> Institutions and statuses need not be created by the speech-act of a single individual; they may be created by the speech-act of a community. Such speech-acts have the form 'We hereby declare it right to greet people known to us' […] Of course, such communal speech-acts are fictitious: we do not create social institutions by speaking in chorus. What happens instead is that the communal institution-creating performative testimony is typically *fragmented and widely distributed over other speech-acts*. The communal performative is never explicitly made; it is only made implicitly or indirectly. It is carried out by people when they do other things: when they talk about greeting their colleagues on the way to work; when they actually

> greet their colleague; when they criticize others for not having greeted them back […] All these other speech-acts – most of which are in fact constatives – 'carry' the relevant communal performative. (Kusch 2002: 67)

In this respect, normatives are claims that either (a) express normative attitudes that are already in place, or (b) are injunctions to or from a community for making sense of things in a certain kind of way. They can be seen as candidate proposals for delimiting the range of meaningful possible ways in which we can understand and make sense of the world.

Why is this significant? A crucial point that I want to draw attention to in this discussion is that rule-governed practices are always, by their very nature, *open*. The open-endedness of practices means that the rules and norms that are in force for a particular practice or community (underpinned by the normative attitudes of participants within that practice or community) are themselves open to challenge, no matter how fixed or stable they might be. They may also be further developed, for example, by 'being extended into areas previously untouched, and refined or resolved for cases where the outcome of their application was unclear' (Peregrin 2014: 86). In that respect, a normative – a claim that somebody or something ought to do or be something – aside from being anchored in the existing practice, also 'reinforces, or challenges, or extends this practice' (ibid: 86). Normatives are not only ways in which we express our commitments, but also ways in which we try to extend and refine those commitments.[7]

The upshot of this section is that norms are clusters of normative attitudes. They are more than regularities in our behaviour. Rather, they depend on the way in which participants in a practice or community adopt normative attitudes to one another in terms of what they say and do. Conceiving of norms in this social, practical manner suggests that a picture of norms as essentially prescriptive (in the sense of dictating what one should do) is problematic. Instead, rules and norms are ways of delimiting the kinds of possibilities for making sense of the world that are available to us. As such, they help to constitute a 'space of meaning'. Furthermore, the restrictive (as opposed to prescriptive) sense of norms is grounded in the open-endedness of our practices, in which the norms governing those practices can be open to challenge and revision. Finally, drawing on Peregrin's analysis, we have seen that normatives (claims or judgements that something is correct or ought to be done) are a key component of constituting and extending the space of meaning in which our practices operate.

[7] Looking ahead, I will argue that scientific models have a similar function with respect to our inferentially articulated scientific understanding. I treat the construction and analysis of models as a tool for articulating, extending and refining our conceptual commitments, by determining the circumstances in which our concepts should apply and the inferential consequences of those applications. In that respect, models are interpreted as having an analogous function to normatives but within scientific contexts.

5 Temporal Normativism and Human Practices

The claim that norms are clusters of normative attitudes thus marks a shift towards patterns of social interaction between participants within a practice. This shift in turn points to an important re-conception of how to understand the notion of a practice itself. For the regulist and regularist accounts, practices are defined as shared sets of rules, behavioural regularities, or some such variant: presuppositions, paradigms, habits, or traditions. On these conceptions, practices are defined by some shared, underlying 'thing' that prescribes how members of that practice or community should act. Treating norms as clusters of normative attitudes offers the resources for an alternative conception.

An initial formulation of this alternative is provided by Barry Barnes (2001). Barnes argues that we should focus not on the idea of a practice as some kind of single object or entity, or that is only made up of coincidentally or statistically convergent sets of individual habits. Rather, we should focus on the interactions between practitioners and their mutual adjustment, coordination and modification of their own actions in response to one another. For Barnes, practices are 'collective accomplishments' of individuals 'concerned all the time to retain coordination and alignment with each other to bring them about' (Barnes 2001: 33):

> What is required to understand a practice of this kind is not individuals oriented primarily by their own habits, nor is it individuals oriented by the same collective object; rather it is human beings oriented to *each other*. Humans beings can ride in formation [Barnes's example of a practice], not because they are independent individuals who possess the same habits, but because they are interdependent social agents, linked by a profound mutual susceptibility, who constantly modify their habituated individual responses as they interact with others, in order to sustain a shared practice. (ibid: 32; citation suppressed)

Barnes's alternative does not rely on either explicit rules or shared regularities and commonalties in behaviour to identify a practice, but rather on the social interaction between practitioners. Understood in this sense, a practice need not depend on any underlying regularity or common performance. Instead, what matters is that agents are responsive to each other in order to achieve the 'collective accomplishment' of bringing that practice to competent execution.

By treating norms as clusters of normative attitudes, the concomitant shift with respect to our conception of practices is to understand practices themselves in normative terms, rather than just as shared regularities. This *normative* conception of human practices has been extensively developed by Rouse, who holds that 'a practice is not a regularity underlying its constituent performances, but a pattern of interaction among them that expresses their mutual

accountability' (Rouse 2007: 48). This conception of practices reflects the idea that norms are constituted by our normative attitudes, in which we take and hold one another to account. On a normative conception of practice, what binds the practitioners together are not explicitly formulated rules, shared regularities in behaviour, or any other thing they mysteriously have in common, nor even a settled, determined goal that all practitioners are directed towards achieving. What matters, on this conception, is that practitioners are held mutually accountable not only to what other practitioners actually do, but to what they *ought* to have done.[8] As Isabelle Peschard formulates this point,

> What is important to qualifying as practitioners in the same practice is not sharing ways of doing, beliefs, or presuppositions, it is being accountable to certain norms, it is one's performances and utterances being subject to questions, to demands of justification, to criticism, to constructive elaboration, being something that matters to the other, something that can make a difference to their own performances and utterances. (Peschard 2007: 148)

Moreover, holding one another to account is itself something that can be done correctly or incorrectly, and this critical aspect can in turn be held to account (Rouse 2007: 3). That is, holding one another to account, treating certain performances as appropriate or inappropriate, sanctioning behaviour, can itself be thought of as normative, and not merely a behavioural disposition to reinforce certain behaviours rather than others (Brandom 1994: 36, 42). For instance, rewarding a piece of behaviour is something that itself can be done correctly or incorrectly. Such a conception of practice reflects a suggestion made by Brandom in an earlier paper: 'we can envisage a situation in which *every* social practice of [a] community has as its generating response a performance which must be in accord with another social practice' (Brandom 1979: 189-90). Such a network of practices need not be identifiable as a regularity of action or belief, however. Consequently, Brandom argues that the difference between regularities and normative responsiveness *is itself a normative issue*, one that concerns the 'distinction between those patterns appropriately explained in causal terms and those appropriately understood as subject to interpretation and normative response' (Rouse 2002: 169). In turn, Rouse suggests that Wittgenstein's remark (noted above), that when we hit justificatory bedrock and all we can say is, 'This is what we do' (*PI*: 217), should not be

[8] A similar precursor to this conception of practice can be found in Davidson (1986) and his well-known claim that 'there is no such thing as a language' (Davidson 1986: 446). We can shed light on Davidson's remarks here if we take them to be directed against regulist and regularist models of natural languages. For Davidson, understanding and using a language involves, 'no learnable common core of consistent behaviour, no shared grammar or rules, [and] no portable interpreting machine set to grind out the meaning of an arbitrary utterance' (ibid: 445). Instead, understanding is an ongoing process of triangulation between the beliefs, actions and meanings attributed to a speaker, the content of what is attributed, and the partially shared circumstances in which attribution takes place.

understood as appealing to a regularity, but instead given the kind of normative inflection with which a parent tells a child, 'We don't hit other children, do we?' Such statements do not describe what children actually do, but function as a response to the child's action, which holds that action to account (Rouse 2007: 4).[9] A normative conception of practice therefore makes normativity an irreducible feature of what it means to be a practice, but does not model this normativity on explicitly formulated rules or on conformity to *de facto* regularities. Instead, what is essential is the capacity of individuals within a practice to adopt normative attitudes to each other.

Thus, on this conception, social interaction has a fundamentally normative character. Norms govern social practices in the sense that individuals within a practice respond to one another and hold each other to account. Normativity is sustained by the interaction between participants of a practice:

> One performance responds to another, for example, by correcting it, drawing inferences from it, translating it, rewarding or punishing its performer, trying to do the same thing in different circumstances, mimicking it, circumventing its effects, and so on. (Rouse 2007: 4)

On this conception, the bounds of a practice are identified by the ways its constitutive performances bear upon one another. To get a handle on this notion, consider chess again. Using chess as an example of a practice does require some qualification. Chess, as we know, involves explicitly codified rules which are taught and learned by participants. However, given the general discussion, we should not see this as a reversion to a regulist model of practices, where a practice is identified by a clearly circumscribed set of rules. On a normative conception, the existence of such rules is not what makes chess a kind of norm-governed practice. What matters instead is that the players and pieces hold each other's actions to account and that their performances be normatively responsive to each other. If I perform a move that violates the rules, such as moving my bishop directly forward rather than diagonally, then to count as a participant in the practice of chess-playing, my competitor should hold my illegal move to account, by calling me out on it, by clarifying what moves are legitimate for a bishop, or by refusing to play if I continue to flout the rule. As we have seen though, whether my competitor holds my transgression to account can itself be done correctly or incorrectly. Of course, not every impropriety is held to account in everyday practice. Nevertheless, if I continue to repeatedly perform illegal moves and my competitor fails to acknowledge and sanction my actions, eventually neither of us could be understood as playing

[9] McDowell (1984: 241-2), Williams (1999: 168-9), and Rietveld (2008: 983) likewise characterise Wittgenstein's 'bedrock' as inherently normative. Thus, the dividing line then between regularist and normative conceptions of practices is how to understand this bedrock: as a complex of mere behavioural inclinations or as inherently normative.

chess any longer. Such interactive responsiveness is what sustains the normativity of a practice like chess.

More significantly, Rouse argues that making normativity an irreducible and defining feature of practices also captures their *temporally extended* character. Practices are not fixed entities, but recurrent patterns of behaviour that evolve over time: 'Practices only exist in continuing to be reproduced. If people stopped producing, exchanging, and consuming goods and services, there would be no economy…if no one taught or undertook courses of study or research, there would be no university' (Rouse 2015a: 193). As such, the normativity of practices needs to account for this. Rouse's proposal builds on the idea that practices are constituted by the mutual interactions of their participants. He argues that what ultimately gives this mutual interaction and accountability its normative force is that there is something 'at issue' and 'at stake' in the outcome of their performances. What is at issue is how and whether the practice should continue and what is at stake in 'the resolution of those issues is what kind of lives the participants can lead, in what circumstances' (Rouse 2016: 559).[10] To illustrate, Rouse draws on Alasdair MacIntyre's conception of a tradition, which both highlights the interactive nature of social practices, but does so in a way that emphasises what is at issue and at stake in a practice:

> What constitutes a tradition is a conflict of interpretations of that tradition, a conflict which itself has a history susceptible of rival interpretations. If I am a Jew, I have to recognise that the tradition of Judaism is partly constituted by a continuous argument over what it means to be a Jew. (MacIntyre 1980: 62)

The point that is drawn from this is that what it is to be a Jew is not fixed by any shared element handed down through tradition, but is rather found in an ongoing concern about what it means to be Jewish. What it is to be a Jew is 'contested among the performances that constitute the ongoing practice of Judaism, in all their historically interrelated complexity' (Rouse 2015a: 165). Because there are no shared rules or fixed underlying presuppositions that antecedently determine what it means to be a member of a particular practice, or even what that practice is, any performance held accountable to a given practice serves as a kind of partial, possibly conflicting, interpretation of what that practice is taken to be and how it should continue. In MacIntyre's example, 'what is at stake among conflicting interpretations of the practice are what it would then mean to be a Jew and to practice Judaism and how those

[10] There are important parallels here between Rouse's normative conception of practices and Heidegger's characterisation of Dasein as the kind of entity for whom its being and existence are at issue for it. Both are characterised by a concern for how and what form their continuation should take. However, spelling out the details of this parallel will have to be the topic of future work.

differences matter. But those differences are not already settled, and there is usually no-agreed-upon formulation of what the issues and stakes are' (ibid: 165).

In this respect, the open-ended nature of a practice is more radical than stated in the previous section. There, the rules and norms that govern a practice were recognised as being open to challenge, revision or extension, the implication being that participants at least agreed on what is at issue or at stake within practice. For Rouse, by contrast, it is quite possible that participants engaged in the same practice may not be in explicit agreement about 'what that practice is, what issues call for revision or repair, or what is at stake in how those issues are resolved' (Rouse 2016: 559). I do not think that this suggests that this is the *de facto* state of our human practices, but instead, that Rouse's account recognises that this is a possibility that is open to any practice. Kuhn's notion of 'crisis science' offers a useful way of illustrating a case when these issues become salient within a practice or system of practices. In periods of crisis, not only are there disagreements between scientists within a given disciplinary community about what issues within that discipline need to be addressed, but more substantially these disagreements extend to the questions of how to address them, what is at stake in resolving them, how that discipline should continue, and so on.

This does not mean that we do not try to express what is at issue or at stake in a practice. Indeed, we often do try to formulate and express through explicit judgements what is at issue or at stake in a given practice. This is, for example, where normative judgements and claims would take on a particular significance. Even attempts to stand outside a given practice should be understood as attempts to articulate and identify the norms that are governing the practice, but such attempts are also accountable to that practice, and in that sense, also form a part of the practice (Rouse 2007: 6). The general point is that any practice, as a temporally extended, recurrent pattern of activity, necessarily outruns any attempt to identify its nature and cannot be reduced to any one specific interpretation.

On the basis of this normative conception of practices, which is grounded in the mutual, social interaction between participants, practices should not be understood as determined by fixed norms, but rather as being subject to ongoing interpretation about precisely what those norms are, what the practice is accountable to, and how it should continue. Norms are not already determinate standards to which performances are accountable but rather 'temporally extended patterns that encompass how we have already been living this part of our lives as well as the possibilities open for its continuation' (Rouse 2015a: 194). What is at issue and at stake cannot be exhaustively determine: in a crucial sense, the promise of resolving what is at issue and at stake motivates the practice to continue, but it is also what the practice aims to clarify. Normativity is an essentially temporally extended phenomenon. As such, normativity is not to be expressed in terms of governance by rules or *de facto* regularities in a community's behaviour, values, or preferences. Normativity is instead sustained through a complex pattern

of interactive relations and clusters of normative attitudes among performances through time, and 'such performances are normative when they are directed toward one another as mutually accountable to common stakes, albeit stakes whose correct formulation is always at issue within the practice' (Rouse 2007: 8).

6 Concluding Remarks

The aim of this chapter was to clarify and expand upon the concepts of norms and normativity. I examined two conceptions of norms - regulism and regularism - and argued that both led to intractable problems. Although the regulist proposal models norms on the kinds of explicit rules that we find in everyday life, I argued that it was inadequate on the grounds that it was subject to a vicious regress of further rules and interpretations. The broadly Wittgensteinian solution to the issue was to find a conception of norms that captured their implicitness within human practices, which is exhibited in our practical behaviour and ability to respond to situations without recourse to explicit interpretations of those situations. The regularist proposal provided an initially plausible alternative to regulism, but it was shown to provide only the semblance of normativity, rather than genuine normativity. Norm-responsive behaviour involves more than just regularities in what we say and do. Instead, norms become expressible through the normative attitudes we adopt towards one another. Norms are clusters of normative attitudes, sustained by the social interactions of participants within a practice.

As clusters of normative attitudes, norms constitute and shape a 'space of meaning', within which what we say and what we do can become candidates for intelligibility and meaningfulness. This space of meaning, embodied and sustained by our norm-governed practices, is open-ended because the norms that govern our practices remain open to challenge, refinement and extension. In this context, I identified normatives as a particular kind of speech act that performs the pragmatic function of making explicit our normative attitudes and thus enabling them to be challenged and revised. This in turn led to Rouse's temporal normativist conception of practices, according to which practices are not defined by fixed sets of norms or sets of shared regularities in behaviour but by patterns of behaviour and performances that are extended over time and mutually accountable to one another. The temporal nature of practices thus embodied a particular form of normativity, which is grounded in what is at issue and what is at stake within a practice. A normative conception thus understands practices as being subject to an ongoing interpretation of what a practice is about and how it should continue.

In the following chapters, I apply the ideas and arguments developed in the previous three chapters to a conception of scientific understanding. My focus in the next two chapters is on

the question of how scientific understanding is facilitated and developed by specific vehicles of understanding: models and explanations. My strategy is to situate the practices of modelling and explanatory discourse within the broader conception of understanding set out so far. Therefore, to conclude this chapter, I shall briefly summarise the main features of this account, which was the focus of the first part of this thesis.

The central claim I have made is that understanding is unintelligible without reference to our norm-governed social practices. Chapter One provided the foundation for this claim in the form of Heidegger's phenomenological conception of what it means to be a subject of understanding. According to Heidegger, understanding is not the product or state of a detached or disembodied mind, but an essential aspect of what it means to be 'in the world'. To be in the world means to be embedded in social, norm-governed practices, paradigmatically in the form of the various practical contexts that define our engagement with the world. More generally, for Heidegger, understanding itself has a dual normative capacity. First, we understand entities in the world on the basis of the possible ways that they can or cannot be, where these possibilities are constituted by general ontological norms of intelligibility. Second, any particular understanding of entities in the world depends upon our self-understanding, and it is this self-understanding that reflects our capacity to be the kinds of subjects that are capable of responding to norms *as* norms.

Chapter Two drew upon Brandom's account of conceptual understanding to address the idea that understanding is constituted by our inferential ability. Although Brandom's account is broadly in agreement with this claim, it departs from other accounts in the current literature on understanding by arguing that conceptual content is only conferred on our inferential abilities in the light of their place within norm-governed social practices that define the 'game of giving and asking for reasons'. Developing a theme from Heidegger, this inferential ability was also seen to be modally robust insofar as it made intelligible counterfactual conceptual possibilities an essential feature of our conceptual understanding. Understanding any concept, for Brandom, involves a grasp of the counterfactual inferential possibilities inherent to that concept. Finally, I argued that Brandom offered a two-dimensional account of conceptual normativity, which is structured by, on the one hand, circumstances in which concepts are applied, and, on the other hand, the consequences of those applications. In turn, working out the inferential links between these two dimensions is imbued with normative significance: determining what should follow from what is a matter of working out which inferential norms we should endorse.

Chapter Three provided a more extensive account of norms and their connection to human practices. The claim that norms are clusters of normative attitudes expands upon the idea that, as subjects of understanding, we are responsive to norms within our practices. Furthermore, the open-ended nature of our norm-governed practices expands upon the (Heideggerian) sense

in which we can respond to norms *as* norms. This was seen not only in the form of those norms being open to challenge and revision, but in Rouse's temporally extended conception of normativity as emerging from what is at issue and at stake in the continuation of our practices. Finally, this broader discussion of rules, norms and practices not only clarifies what is meant by the claim that understanding is unintelligible without reference to our norm-governed practices, but also extends our conception of what is involved in understanding at all. Thus, Rouse writes:

> The performances that constitute conceptually articulated practices are both socially differentiated and dynamically responsive to that differentiation through ongoing efforts to sustain the coherence of a common discursive practice. Conceptual understanding is then not the grasp of a static holistic structure but an active capacity to track, adjudicate, and respond appropriately to the more or less divergent performances within social practices […] (Rouse 2015a: 83)

Thus, our understanding reflects not just a particular cognitive attitude or relation to a body of knowledge, but the way in which we are situated within norm-governed social practices and, indeed, our general orientation to the world. Our understanding is a dynamic, active and socially mediated capacity to make sense of the world.

Chapter Four

The Meaning of Models: Model-Based Understanding and Conceptual Articulation

1 Introduction

Models are a ubiquitous and indispensable feature of scientific research. Recent work on their construction, use and analysis has been largely driven by the growing interest in scientific practice, as opposed to formal reconstructions of scientific knowledge. Historians and sociologists of science have amassed a great number of detailed case studies that describe and analyse scientific models and their function within research practices.

This chapter examines how models facilitate scientific understanding. A central epistemic issue for accounts of model-based understanding, which I discuss below, is that the use of models within scientific research challenges the idea that understanding is factive in the same way that knowledge is (i.e. if S knows that *p*, then *p* is true). The problem is that models are typically constructed in ways that either deliberately misrepresent the world or use idealizations that are, strictly speaking, false. This provokes the question of how idealized or unrealistic models could contribute to scientific understanding of how the world actually *is*. Such a concern is exacerbated by the fact that some of the dominant accounts of scientific understanding, such as de Regt's or Grimm's, say little about how understanding is provided by idealized models.

In this chapter, I dispute the idea that models only provide understanding in virtue of being factive or representationally accurate. This is not to deny that our understanding should be accountable to the world in some sense, but to take this as the primary epistemic worry misconstrues the role models play in providing scientific understanding. The issue I suggest instead concerns conceptual meaning and significance, rather than truth or falsity. Specifically, I argue that models afford scientific understanding through their role in *conceptual articulation*. Conceptual articulation with models consists in working out the inferential consequences of our concepts, what they mean, and where and how they apply to the world. To clarify the notion of conceptual articulation as a constitutive element of model-

based understanding, I focus on three related aspects of it: (i) its intimate connection to the *intelligibility* of a domain of entities: scientists use models to determine what can be meaningfully said about entities within that domain; (ii) its contribution to a *modal* understanding of the target phenomena: the way in which models afford understanding of what is possible or not possible for a domain of entities; and (iii) the way in which conceptual articulation is bound up with conceptual normativity, an idea which concerns which inferences *ought* to be endorsed by a scientific community. In this chapter, the central claim is that models provide understanding by enriching what scientists can meaningfully say about the world.

2 Models: An Overview

It is customary to refer to the model as the 'source' and the thing being modelled as the 'target'. The target is the phenomenon or phenomena that the scientist is trying to make sense of.[1] This terminology should be treated minimally at this stage; it implies no commitments to what types of things the source and target can be: they may be concrete or abstract, physical or mathematical, real or imaginary (Suárez 2015: 41). Models come in different types. In *Models and Metaphors*, Max Black (1962) distinguished between four overarching types of models found in science: scale models, analogue models, mathematical models, and theoretical model.[2]

Scale models enlarge or shrink certain features of their target system, such as a globe as a model of Earth (Bailer-Jones 2009: 3). Often, scale models are used to simulate the behaviour of their target systems, without seeking to explain why it behaves as it does. The San Francisco Bay Delta Model would be an example (Weisberg 2013). *Analogue models* function in a similar way, but involve a change of medium. The Phillips machine, for example, represents the workings of the macro-economy by the ebb and flow of coloured water in a hydraulic system (Morgan and Boumanns 2004). In turn, *mathematical models* make use of mathematical language and formalism to represent their target system, for example, the classical simple harmonic oscillator model of the mass-spring system. Finally, *theoretical models* frequently employ abstract concepts, idealizations, and theoretical principles. Familiar examples here include the billiard ball model of gases, Bohr's model of the atom, the liquid-drop model of the nucleus, or the Lotka-Volterra model of predator-prey interaction. The use

[1] Following Bogen and Woodward (1988), 'phenomena' refer to stable regularities in the world, as distinct from individual instances of observational or experimental data. For example, lead melting at 327.5 ˚C (phenomenon) as distinct from a series of temperature readings as a piece of lead is heated up (data).

[2] This list is by no means exhaustive. I only offer it here as a useful entry point to the variety of types of models in scientific practice. See Frigg (2012) for a more expansive discussion.

of 'theoretical' to describe these models is intended in a reasonably broad sense and includes any case in which scientists deliberately simplify or idealise a system or phenomenon to explain or predict an aspect of its behaviour (Toon 2012: 9).

In recent years, philosophical interest in models has been driven by close attention to how models are used in scientific practice.[3] An influential idea in this context is that models are partially autonomous *mediators* between theory and reality (Morgan and Morrison 1999). On this view, models are to be regarded neither as a merely auxiliary intermediate step in applying scientific theories, nor as constructed purely from data. It is argued that this is because it is argued that there is no algorithm for constructing adequate models from theoretical principles alone. Thus, more recent work typically rejects what Cartwright (1999) labels the 'vending machine' view of theories and models: the view that models are neatly contained within a theory and all that is required is the right kind of input in order to derive a particular model. Instead, it is now widely acknowledged that scientific models are often constructed from a heterogeneous mixture of elements, such as simplifications, metaphors, policy views, empirical facts, as well as theoretical ideas. As Marcel Boumans colourfully put it, model building can be 'like baking a cake without a recipe' (Boumans 1999: 67). This is not to suggest that theories do not guide model construction; merely that that the construction of models, especially new models, is more like a trial-and-error process of working out how these various elements fit together.

Michael Weisberg (2007) and Peter Godfrey-Smith (2006) have proposed that model-based reasoning in science can be distinguished from other theoretical activities through its *indirect* nature of representation. Weisberg suggests that modelling proceeds in three stages. First, a model is constructed, then, second, the modeller refines, analyses and articulates its properties and dynamics. It is not until the third stage that the relationship between the model and its target is assessed, 'if such assessment is necessary' (Weisberg 2007: 209). By calling modelling 'indirect', Weisberg and Godfrey-Smith draw attention to the fact that models are not primarily constructed to represent their target systems as faithfully as possible. Rather than studying real-world phenomena directly, modellers construct and study idealized surrogate

[3] Typically, this move is framed as a reaction against both 'traditional' syntactic and semantic accounts of theories, from which earlier accounts of models were derived. On the syntactic version, a theory is a set of claims and models of the theory, such that all the claims of the theory are satisfied and no additional assumptions are imported except those which are legitimately grounded in a description of the phenomena to be represented. Semantic accounts held that theories *are* sets of models (e.g., Van Fraassen 1980; Giere 1988). The syntactic view has been widely discredited, but the semantic view is still defended (e.g., French and Ladyman 1999; da Costa and French 2000).

systems to which only a few select properties are attributed.[4] The important insight of the notion of indirect representation is 'to redirect the focus from models to the activity of modelling' (Knuuttila and Loettgers 2017: 1010).

As an example, the Lotka-Volterra model of predator-prey interaction was constructed to study the dynamics between two idealized populations of organisms. Unlike real populations of organisms, the kind of properties stipulated to be possessed by these model populations were very few, such as intrinsic exponential growth rate for the prey in the absence of predators and a constant death rate for the predators (Weisberg 2007: 210). Even so, the model was used to infer empirically verifiable conclusions concerning the dynamics of fish populations in the Adriatic after the First World War. Often, the motivation for constructing a model is to reduce the amount of complexity that must be accounted for. With an idealized model system to hand, scientists can focus on the effects of a few choice properties or mechanisms. By studying how variations in the assumptions out of which the model is built lead to different or similar outcomes, scientists can infer conclusions about its real-world counterpart (the 'target system').[5] Thus, modelling is a form of surrogative reasoning, which 'designate[s] those cases in which someone uses one object, the *vehicle* of representation, to learn about some other object, the *target* of representation' (Contessa 2007: 51).

It has been argued that the understanding of models as surrogate systems is analogous to the epistemic dynamics of models and experiments (Mäki 2005; Morgan 2003). According to Mäki, modelling and experimentation are both attempts at *isolating* the casually relevant factors. Whereas in the case of experimentation, such isolation is achieved causally through experimental controls, in the case of models, such isolation is achieved by making more or less unrealistic assumptions (Kuorikoski and Lehtinen 2009: 121). The unrealistic nature of such assumptions derives from the use of *idealizations* and *abstractions* in the construction of models. As several philosophers of science have argued (e.g., Cartwright 1983), scientific hypotheses typically do not apply, strictly speaking, to any real system in the physical world, but rather to idealized versions of the phenomena. As Roman Frigg puts it:

> When studying the orbit of a planet we take both the planet and the sun to be spinning perfect spheres with homogenous mass distributions gravitationally interacting with each other but nothing else in the universe […] [W]hen studying the exchange of goods in an economy we

[4] The notion of models as 'surrogate systems' has been discussed in Swoyer (1991) and Mäki (2005).

[5] Weisberg contrasts this with 'abstract direct representation' which refers to the analysis of real-world phenomena (abstracted and represented in certain ways) without any mediating model. For example, Weisberg argues that Mendeleev's ordering of chemical elements is a case of abstract direct representation, since Mendeleev worked directly with representations of real-world phenomena, the chemical elements, rather than through the construction and analysis of an idealized model of those elements (Weisberg 2007: 216).

> consider a situation in which there are only two goods, two perfectly rational agents, no restrictions on available information, no transaction costs, no money, and dealings are done in no time. (Frigg 2010: 251-252)

Idealizations are typically introduced into models as conscious misrepresentations of some aspects of a target system in order to render other factors of interest more salient.[6] The billiard ball model of gases, for instance, treats the molecules of a gas as composed of dimensionless, spherical molecules that are not acted upon by gravity and do not act upon each other in collisions. Of course, no such gas exists, but treating gases in this way allows scientists to selectively focus on specific aspects of a phenomenon, while ignoring others, and to make such aspects computationally tractable. Other examples include frictionless planes, Newtonian point masses, and, in the case of the Hodgkin-Huxley model of action potential, the representation of neurons as electrical circuits.

Both idealization and abstraction are an essential part of building a model. This is because the variables exerting some influence on a phenomenon of interest are often too numerous to take into account. Idealizations and abstractions allow the system to become more analytically tractable and computable. For instance, no real system strictly obeys the laws of physics, so using these laws to explain the behaviour of real systems requires the use of various approximations, abstractions and idealizations. By idealising selected parts of a phenomenon, scientists can draw attention to and home in on those aspects they take to be relevant to the problem at hand, while ignoring other features. In doing so, scientists make choices about which idealising assumptions should be involved in the model. The nature of these purposes and epistemic goals are dictates how the model is used and which aspects of the target system are to be focused on.

Hence, as partially autonomous surrogate systems, models provide scientists with a range of tools and can be used for a variety of different purposes. As well as enabling scientists to infer conclusions about their real-world counterparts, models are often employed for heuristic or predictive purposes. In these ways, models are used for determining which aspects of a target phenomenon to focus on, which concepts can be brought to bear on those phenomena, and what the implications of modelling the target in a specific way would be. Because of this, recent philosophical interest in modelling practice is often less concerned with what models *are* and more concerned with what models are *used for*, i.e. their various functions and purposes in scientific research (e.g., Keller 2000).

---

[6] Two types of idealization are distinguished in the literature on scientific models, namely Aristotelian and Galilean idealizations. Aristotelian idealization refers to the process of stripping away properties of the target that are taken to be irrelevant. This kind of idealization is often called 'abstraction'. Galilean idealization refers to the conscious and deliberate misrepresentation of a target system in order to make it more tractable for study.

3 Understanding with Models

One of the key philosophical issues facing model-based understanding concerns factivity. The issue is that, as Grimm (2006) observes, philosophers of science tend to agree that scientific understanding is a species of knowledge, and hence factive.[7] To say that knowledge is factive is to say that, if *S* knows *p*, then *p* is true. Consequently, if *S* understands *p*, then *p* must also be true. Furthermore, if, as some philosophers have argued (e.g., Khalifa 2013; Strevens 2013), all scientific understanding reduces to explanatory understanding, and if understanding results only from explanations based on theories that are (at least approximately) true, then this also would imply scientific understanding must be factive. By contrast, epistemologists have tended to disagree that understanding must be factive in all cases. It is generally accepted in epistemology that there are different kinds of understanding. For example, objectual understanding – which involves having a grasp of the connections between items of information in a body of knowledge – is held to be able to tolerate some falsehoods. A few falsehoods at the margins of one's objectual understanding might degrade it, but do not undermine it completely (Baumberger 2011: 81). 'Moderate factivists' (for example, Kvanvig 2009, Mizrahi 2012) hold that objectual understanding remains possible even with some false propositions, if the core central propositions remain true.

Some philosophers have argued, however, that factivity is far too demanding and restrictive for an adequate conception of scientific understanding. According to Elgin, a factive conception of understanding 'does not reflect our practices in ascribing understanding and it forces us to deny that contemporary science embodies an understanding of the phenomena it bears on' (Elgin 2007: 33). Similarly, de Regt (2015) argues that the traditional view of understanding 'should be replaced by an alternative conception that allows for understanding without truth' (de Regt 2015: 3794-3795). A key argument for both Elgin and de Regt concerns the use of models in the history and practice of science, and specifically the fact that those models almost invariably incorporate idealized assumptions, fictions and deliberate falsehoods, which do not mirror how the world is. These idealizations are thus, strictly speaking, false, but they 'can neither be eliminated from scientific theories nor banished to their periphery' (Baumberger 2011). Thus, a factive conception of understanding is held to be inadequate because the use of idealizations in scientific theorising and modelling is a central and ineliminable feature of scientific research. Elgin uses the Ideal Gas Law as her primary example:

[7] See, e.g., Achinstein (1983) and Trout (2002).

> Science streamlines and simplifies. It devises and deploys comparatively austere models that diverge from the phenomena it seeks to explain. The ideal gas law accounts for the behaviour of gases by describing the behaviour of a gas composed of dimensionless, spherical molecules that are not subject to friction and exhibit no intermolecular attraction. There is no such gas. Indeed, there could be no such gas. Nonetheless, scientists purport to understand the behaviour of actual gases by reference to the ideal gas law. (Elgin 2007: 38)

Elgin's point is that the use of idealized and unrealistic models in scientific practice is compatible with a significant degree of understanding. For Elgin, to deny that idealized models can yield understanding would force one to deny understanding to much of contemporary and historical scientific practice. This, she thinks, would be substantially at odds with the clear cognitive success that the sciences enable.

De Regt, in turn, draws on a range of examples from the life and social sciences to show that the use of idealized and unrealistic models is also compatible with advances in scientific understanding. Economic models, for instance, are highly abstract and idealized, entailing regularities that cannot be found in the real world. Much of classical economics, for example, is founded on the conception of human beings as rational utility maximizers, even if we know from psychology and behavioural economics that this is rarely the case. Nevertheless, it would be controversial to suggest that classical economics provides *no* understanding of economic reality. As such, de Regt argues that 'unrealistic models provide understanding that is used in explanations of complex economic reality' (de Regt 2015: 3787). In the life sciences, the use of 'model organisms' (for example, the bacteria *E.coli*, the fruit fly *Drosophila melanogaster*, and the mouse cress *Arabidopsis thaliana*) to understand biological processes are unrealistic in the sense that they idealize and abstract from biological reality. Although model organisms are real organisms, they differ in important respects from organisms 'in the wild'.[8] First, model organisms are not typically selected for their representative applicability, but for pragmatic reasons, such as experimental manipulability and tractability, breeding time, and costs. Second, they are standardised, which ends with them often being 'extremely distant from those that could be easily found outside the laboratory and in nature' (Ankeny 2009: 200-201). Thus, even model organisms involve idealizations which are used to make sense of complex biological reality, while not representing it faithfully (de Regt 2015: 3789).

Given that the models and idealizations they appeal to are drawn from examples central to a discipline's understanding of the phenomena they target, Elgin and de Regt conclude that the cognitive success that we attribute to scientists does not involve attributing success in the sense of knowledge about the truth with respect to a particular domain in nature. What their examples suggest is that scientists will often forego representational accuracy or factive

[8] For discussion of model organisms, see Ankeny (2009) and Ankeny and Leonelli (2011).

understanding in favour of computational power or experimental tractability. As such, 'an approximately true description of the system is no precondition for understanding' (ibid: 3789). I am sympathetic to Elgin's and de Regt's conclusions; I am inclined to agree with them that a factivity condition is too restrictive and demanding for an adequate conception of model-based scientific understanding. However, I do so for different reasons. To see why, let us pause for a moment to consider one of the objections to Elgin and de Regt's conclusions.

The objection comes from Kvanvig (2009: 342-343) and Mizrahi (2012: 246-249) in defence of moderate factivity. Kvanvig argues that using an idealized model to understand some target phenomenon requires that one knows what the model is like and how it relates to the target. This involves knowing the extent to which the model is an idealization and how this distorts or is unrealistic with respect to the target phenomenon. If a scientist knows all this, Kvanvig contends, then the factivity requirement for understanding is not violated. Similarly, Mizrahi responds directly to Elgin's use of the Ideal Gas Law, and argues that we need to distinguish between the Ideal Gas Law, the conditions under which it applies, and the idealizing assumptions necessary to derive the Ideal Gas Law. Mizrahi contends that the idealizing assumptions in the Ideal Gas Law only lie at the periphery of the model, and as such, the fact that they are false does not detract from the central propositions of the gas laws, which must be true if scientific understanding is to be obtained, thus concluding that moderate factivity is preserved (Mizrahi 2012: 250).

Kvanvig's and Mizrahi's responses to Elgin and de Regt presume that scientists know what the idealizing assumptions say about the world, how they apply, and the extent to which they diverge from reality. It is thus because of this knowledge that we can say idealized models yield understanding. The difficulty with this claim, however, as Baumberger has pointed out, is that in some cases scientists do not know how their models diverge from reality, or under which conditions exactly that their model gets it right. Nonetheless, this does not seem to undermine their understanding of the target phenomenon. Environmental scientists, for example, do not know exactly how their climate models diverge from reality, but this does not completely undermine their understanding of global warming (Baumberger 2011: 83).

Baumberger treats this point as indicating how understanding and knowledge can come apart. According to Baumberger, Kvanvig and Mizrahi do not succeed in showing that knowledge – and by extension, truth or factivity – is necessary for understanding. However, I think this point reveals something different. What Baumberger, Elgin and de Regt are highlighting, in my view, is the distinction between conceptual meaning and truth. On this interpretation, what Elgin's and de Regt's arguments against factive conceptions of model-based understanding point to is the fact that, while reference to truth and factivity can mark out one way we can evaluate model-based understanding, we need to be sensitive to the fact that conceptual meaning and truth come apart. This points to a more specific elaboration of

my discussion of Brandom's two-dimensional model of the normativity of conceptual understanding. There (Chapter Two, section six), we saw that the normative accountability of our conceptual understanding spans two dimensions. On the one hand, what we say and do can be assessed in terms of the meaningfulness of our claims and actions (whether what we say or do *makes sense*), and on the other hand, we can determine whether what we say or do is *justifiable* or *true*. As Ian Hacking once noted, '[w]hether a proposition is as it were up for grabs, as a candidate for being true-or-false, depends on whether we have ways to reason about it' (Hacking 2002: 160). The distinction between these two distinct but interconnected poles of normativity means that prior to determining whether what we say about the world is true or justified, it must already be understood at the level of conceptual meaning and intelligibility.

From this perspective, Baumberger is right to say that scientists' understanding is not undermined if they do not know how their models diverge from reality, or under which conditions the model correctly applies. The possibility of settling these issues presupposes that the models and their target phenomena already make sense and are understood. With regards to models and their use in scientific practice, the issue then is making sure not to conflate meaning with truth. Similarly, what Elgin's and de Regt's examples reveal is that we can judge whether models provide understanding – whether they enable us to grasp and make sense of the world – without having to conflate that with the issue of whether what models say something that is true about the world.[9] Put more generally, the point is that this distinction between meaning and truth enables us to distinguish between how we *take* the world to be from how it actually *is*.[10] The point is summarised by Rouse: 'We cannot ask whether a theory, a law, or any other hypothesis is true unless we have some understanding of what it says, and to which circumstances it appropriately applies' (Rouse 2009: 38).

Therefore, my aim in this chapter is to focus on the way which models provide understanding at the level of conceptual meaning, intelligibility and significance. To do this, I argue that models afford scientific understanding through their role in *conceptual articulation*. Conceptual articulation with models consists in working out the inferential consequences of our concepts, what they mean, the circumstances to which they appropriately

[9] I am not denying here that the truth or justification of what models enable us to say is important to conceptual understanding – it obviously constitutes one pole of the normativity of our conceptual understanding. What I want to emphasise is that while we will evaluate what models say about the world along both poles in practice, it is important not to conflate meaning with truth within our analysis. Ultimately, both conceptual meaning and truth are essential for scientific understanding.

[10] Importantly, neither Elgin nor de Regt deny some role for truth in scientific understanding (e.g., de Regt 2015: 3795). They recognise and grant that scientific understanding must be, in Elgin's words, 'tethered' to the world (Elgin 2007: 35) – a point which I agree with. Their point is that focusing solely on the way in which models enable factive understanding risks obscuring the contributions modelling in scientific practice makes to conceptual meaning and intelligibility.

apply and the consequences of their application. I begin in the following section by clarifying the notion of conceptual articulation in more detail before returning to models in section five.

4 Scientific Research as Conceptual Articulation

4.1 Kuhn on Paradigm Articulation

An early formulation of conceptual articulation can be found in Thomas Kuhn's *Structure of Scientific Revolutions*. Recall that for Kuhn normal scientific research takes place against the background of a paradigm, which provides a set of values, practices and basic concepts that scientists hold in common. Paradigms serve to guide further scientific work and 'normal science' takes place within a paradigm once it has been established. According to Kuhn, there are three types of normal scientific research projects that take place against the background of intelligibility provided by the paradigm (Kuhn 1996: 25-30). The distinction between these three types of normal science research is an analytic one, whereas in practice, 'they are neither always nor permanently distinct' (ibid: 25).

First, a paradigm highlights significant classes of facts which require further investigation and determination. Typically, the scientific work involved here is geared towards providing more precise and specific quantitative determinations of those facts so that they can be deployed in solving other outstanding problems. A second aspect of scientific research concerns testing the paradigm empirically, to see how it measures up against nature. This should not be understood in the Popperian manner of a project of wholesale falsification, but more in a 'diagnostic' sense which helps to identify whether, where, and how the paradigm needs further articulation and refinement (Rouse 2003: 110). The third and final type of research project is paradigm articulation. Kuhn's examples of this include the determination of universal constants, the development of quantitative laws, and the selection of ways to apply the paradigm to a related area of interest. Thus, paradigm articulation is the process of working out the conceptual understanding supplied by the paradigm by showing, in finer-grained and more empirically tractable detail, how its concepts apply to situations in the world. By doing so, the sciences and their research practices extend and refine 'our conceptual grip on significant aspects of the world' (Rouse 2013: 62).

Key to the notion of conceptual articulation is the claim that conceptual content is open-ended or open-textured, rather than fully determinate. That is, conceptual articulation presupposes the idea that what our concepts say about the world and what they mean is never

fully worked-out, but always subject to ongoing extension and revision.[11] This idea is nicely captured by recent perspectives on scientific practice that emphasise the epistemic uncertainty and conceptual openness of much scientific research. For example, the prospective uncertainty at the heart of the research process is central to Hans-Jörg Rheinberger's framework for analysing what he calls 'experimental systems' (1997). Rheinberger argues that when scientists investigate a given object of research, this object is, by definition, ill-understood. He refers to such objects as 'epistemic things' to emphasise that they are 'things' towards which scientists direct their epistemic attention without having a very clear idea of what their important features are.

A useful illustration of the open-ended character of scientific concepts, which forms part of the context for what I argue about model-based understanding, is Kuhn's account of thought experiments (Kuhn 1979).[12] For Kuhn, thought experiments become important when scientists 'have acquired a variety of experience which could not be assimilated by their traditional mode of dealing with the world' (Kuhn 1979: 264). Further, Kuhn suggested that thought experiments were like real experiments insofar as they articulate concepts by working out their application to novel situations (ibid: 242). Kuhn argued that the goal of extending concepts through thought experiments is to elicit conceptual revision of some kind. However, for Kuhn this does not simply mean that the function of thought experiments is to find conceptual confusions already implicit in our familiar uses of concepts. Rather, thought experiments work by extending scientific concepts *beyond* our familiar uses so that they come into conflict or tension in the imaginatively extended situation. Although concepts acquire meaning and sense in their application in familiar concrete settings, Kuhn thought that their meaning cannot be confined to those settings. By working out how to apply concepts in new, unforeseen circumstances, thought experiments provide ways of exploring the intelligibility of our concepts, their inferential relations, and what their uses commit us to.

This process is not just a case of internal conceptual clarification and organisation (ibid: 242). Rather, Kuhn argued that in the process we learn something about the world itself to the extent that the uses of our concepts can be applied intelligibly to the world. By extending concepts beyond their familiar application and finding that they do not fit situations to which they should apply, thought experiments can prompt conceptual revision, which in turn can teach us something about the world and not only about our conceptual apparatus. As such, Kuhn identified two conditions on the proper functioning of thought experiments. First, thought experiments must draw on the familiar uses and applications of concepts, and employ

[11] This idea builds on work by Mark Wilson (2006), Joseph Rouse (2015a), and Steinle and Feest (2012).

[12] On the relation between thought experiments and scientific understanding, see Stuart (2016).

them in ways that do not strain normal usage: 'the imagined situation must allow the scientist to employ his usual concepts in the way he has employed them before' (ibid: 264-265). The aim of extending scientific concepts beyond their familiar uses is to bring about a conflict or tension in the imaginatively extended situation. Second, 'though the imagined situation need not be even potentially realizable in nature, the conflict deduced from it must be one that nature itself could present; indeed […] it must be one that, however unclearly seen, has confronted him before' (ibid: 265). Thus, Kuhn states that:

> The concepts "corrected" in the aftermath of thought-experiments displayed no intrinsic confusion. If their use raised problems for the scientist, those problems were like the ones to which the use of any experimentally based law or theory would expose them. They arose, that is, not from his mental equipment alone but from difficulties discovered in the attempt to fit that equipment to previously unassimilated experience. Nature rather than logic alone was responsible for the apparent confusion. (ibid: 261)

Therefore, Kuhn argues that by using thought experiments, we could learn something about the world insofar as it resists certain applications of concepts. In his primary example, the difference between instantaneous and average velocity only becomes conceptually salient in circumstances where comparisons of velocities in those terms diverge (ibid: 249-251). Against the background of these conditions – the prior employment of concepts and intelligible possibilities in nature – thought experiments can 'articulate concepts by presenting concrete situations that display differences that are intelligibly connected to prior understanding' (Rouse 2009: 40). But thought experiments can only function like this, Kuhn argues, because the concepts they employ are open-textured, rather than fully determinate. By working out how to apply concepts in new, unforeseen circumstances, thought experiments can retroactively transform their use in more familiar contexts. For Kuhn, conceptual articulation forms a substantial part of ongoing research practices and thought experiments are one type of method for undertaking this. On this view, the purpose of conceptual articulation is to work out the inferential commitments embedded in the conceptual understanding supplied by the paradigm.

4.2 *Rouse on Conceptual Articulation*

The Kuhnian emphasis on conceptual articulation as an essential feature of scientific research has been developed in more recent years by Rouse. For Rouse, conceptual articulation forms a fundamental part of how the sciences investigate the world. To see this, I shall highlight three aspects of Rouse's account of conceptual articulation and how they extend the basic Kuhnian idea.

First, Rouse joins Kuhn in rejecting the idea that conceptual content can be fully determined. The idea of fully-determinate conceptual content is perhaps most obviously associated with the 'classical' picture of concepts, where the meaning of concepts is analysed in terms of necessary and sufficient conditions, but the association also extends to causal and descriptive accounts of concepts. For both, the meanings of concepts are fixed, either through being referentially and causally connected to a state of affairs in the world, or through their place within a background theory or network. For Kuhn and Rouse, by contrast, conceptual meaning and understanding is never fixed, but always subject to ongoing extension and revision. Conceptual articulation can thus be thought of as the process of working out more precisely what to do with those concepts in our understanding, but without the assumption that what is worked out will remain stable or fixed. As Rouse argues, 'The concepts employed are understood as open textured in ways that both permit and encourage further articulation or correction of previous patterns of use. What scientific claims say about the world is thereby always open to further transformation' (Rouse 2015a: 212). This idea is illustrated by Brandom, who draws an analogy between conceptual understanding and grasping a stick. In grasping a concept, like grasping a stick, we only take hold of it from one end, but nevertheless are accountable for our use of the entire domain of that concept, even if we may not fully understand, or grasp, each part of it and its consequences (Brandom 1994: 583). The underlying point here is to recognise that concepts often commit us to more than we know how to say or do, and before we can work out whether we are entitled to the use of a concept, we need to work out what it says about the world. This is not to imply that understanding a concept means grasping what it means and how it applies in every possible, relevant situation (Rouse 2015a: 238). Rather, it is calling attention to the fact that our conceptual commitments often outrun our more limited understanding of how those concepts would apply in different and unfamiliar circumstances. On this view, a crucial component of scientific research is the articulation of our concepts by working out how and where they apply.

The second aspect of Rouse's approach to conceptual articulation is his rejection of the assumption that this process must be a wholly internal, intralinguistic or mathematical matter of developing and regulating inferential relations, such as we find in Quine's image of a 'web of belief'. According to Quine, scientific theory forms a self-enclosed field of beliefs about the world that only encounters experience at its periphery. On this model, conceptual development is an entirely theoretical affair, a matter of regulating inferential relations among sentences within the web of belief. To the extent that Quine does recognise the contributions of experimental activity, it consists in the form of experiences that impinge on our web of beliefs from the outside, which provide occasions for conceptual development. By contrast, Rouse argues that attention to scientific - and particularly experimental - practices undermines the idea that conceptual articulation and development is merely intralinguistic. Instead, he

claims, experimental practices, and the establishment of experimental systems, substantially contribute to the articulation of conceptual understanding.

There has been considerable philosophical attention on experiment and material practice over the past several decades, which has to a large extent discredited the idea that the function of experimentation is primarily to test theories or to decide between hypotheses.[13] Ian Hacking challenged Kuhn's account of normal science by pointing out that Kuhn assumed that the aim of normal research is to fill out an existing theoretical perspective. Richard Burian (1997) and Friedrich Steinle (1997) developed the idea that scientific practice includes 'exploratory research' aimed towards uncovering new information that is not closely related to any existing theory or theoretical perspective. Moreover, as Uljana Feest argues, there is an inherent conceptual openness to experimental work, since 'experiments often serve the function of exploring a given subject matter within an as of yet conceptually open territory' (Feest 2016: 41).

The third and final aspect I want to consider is Rouse's more general rejection of the idea that the goal of conceptual articulation in the sciences is to produce a complete, systematic or unified picture or representation of the world. Instead, Rouse argues that what matters first and foremost is that the sciences render aspects of the world intelligible, in such a way that we can investigate, discuss, and reason about them. Before we can ask whether our concepts and the theories and models that employ them do adequately represent the world, or whether and to what extent they are justified, *we must already have some understanding of what they say about the world*. Questions of the representational accuracy of scientific hypotheses, models and claims presuppose that we already have some understanding of the intelligibility of those claims. To articulate our conceptual understanding is thus to refine and develop what that understanding says about the world.

A primary motivation for this view comes from Sellars.[14] In 'Empiricism and the Philosophy of Mind', Sellars famously argued that, 'in characterizing an episode or a state as that of knowing, we are not giving an empirical description of that episode or state; we are placing it in the logical space of reasons, of justifying and being able to justify what one says' (Sellars 1997: 76). Rouse's account of conceptual articulation draws upon and develops Sellars' image, arguing that, rather than identifying scientific understanding with some position *within* the space of reasons as a body of knowledge, scientific research should instead

[13] For instance, in Hacking (1983), Franklin (1989), Radder (1996), Galison (1987, 1997), Rheinberger (1997), and Chang (2004, 2012).

[14] The idea is also latent in Heidegger's account of scientific research. For Heidegger, a crucial component of scientific research lies in '…the articulation of the understanding of being, the delimitation of an area of subject-matter (a delimitation guided by this understanding), and the sketching-out of the way of conceiving which is appropriate to such entities' (*BT*: 414).

be conceived of as an ongoing reconfiguration of the entire conceptual 'space'. Scientific research does not just presuppose a stable background provided by the space of reasons, but continually develops and configures that space. Thus, on this view, the function of scientific research is to,

> […] bring aspects of the world into the space of reasons by articulating them conceptually. This achievement allows them to be recognized, discussed, understood, and responded to in ways open to reasoned assessment. The sciences also revise the terms and inferential relations through which we understand the world, which aspects of the world are salient and significant within that understanding, and how they matter to that overall understanding. Such revision is ongoing and built into how we articulate the world conceptually. (Rouse 2016b: 29-30)

Rouse holds that scientific research practices are essential to this process for it is through the introduction of new experimental systems, practices, and skills, as well as developing new models and other theoretical formulations, that sciences open new possibilities for conceptual articulation and continually reconfigure the scientific space of reasons (Rouse 2015a: 26). Furthermore, conceptual articulation does not begin *de novo*; rather, scientists are always already embedded in research traditions, communities, and practices that provide a prior understanding that guides inquiry, even if that understanding is subsequently extended substantially or rejected altogether. As I shall argue, models offer new opportunities to extend, develop and refine our conceptual understanding.

In this section I have clarified the notion of conceptual articulation with reference to Kuhn and Rouse. I highlighted three aspects: (i) that conceptual articulation presupposes the open-ended or open-textured nature of scientific conceptual content; (ii) that conceptual articulation does not have to be restricted to a process of verbal, linguistic or mathematical investigation, but can, and often does, incorporate our material practices too; and (iii) that conceptual articulation is not merely a matter of providing a more systematic or finer-grained representation of the world. Rather, it involves an ongoing configuration of the space of reasons out of which claims concerning the representational status of models and the norms that underpin them become intelligible. My contention is that models and modelling practice form a crucial component of this process of conceptual articulation.

5 Models as Tools for Conceptual Articulation

My claim is that models provide understanding through their contribution to the process of conceptual articulation. In arguing for the truth of this claim, I combine an inferentialist

conception of models with a view of models as epistemic tools.[15] On this view, models function as *tools for inferential reasoning.* In this section, I highlight three ways in which models as inferential tools facilitate conceptual articulation: (1) their role in constituting the *intelligibility* of a domain of entities, which refers to the use of models to determine what can be meaningfully said about entities within that domain; (2) their contribution to a *modal* understanding of the target phenomena: the way in which models afford understanding of what is possible or not possible for a domain of entities; and (3) their bearing upon which inferential patterns *ought* to be endorsed by a scientific community.

5.1 Models as Inferential Tools

My approach is functional: I am concerned with the uses and functions of models in scientific practice. The idea of analysing models in terms of their function within scientific inquiry has been developed more fully in recent work on models in the engineering sciences by Tarja Knuuttila and Mieke Boon.[16] Building on these studies of modelling practice, where the ability to intervene in phenomena, rather than simply representing them, is more important, Boon and Knuuttila argue that it is best to 'consider scientific models in engineering as *epistemic tools* for creating or optimizing concrete devices and materials' (Boon and Knuuttila 2008: 695). This, they argue, extends to modelling across the sciences. In so doing, they build on the pragmatist insight that models are constructed for particular epistemic goals and purposes, but argue for a more robust characterisation of the epistemic features of models that allow them to achieve these goals. What is central to their approach is the emphasis on the *concrete* dimension of models. Models ought to be regarded 'as concrete artefacts that are built by specific representational means and are constrained by their design in such a way that they facilitate the study of certain scientific questions […] by means of construction and manipulation' (Knuuttila 2011: 262).

It might be argued that this emphasis on the concrete dimension of modelling is limiting. It would be difficult to say, for instance, in what sense abstract or mathematical models can be subject to this analysis. Consequently, the epistemic tools approach only applies to a restricted class of models. However, there are two replies to this concern. On the one hand, Boon and Knuuttila argue that even in such cases, model building still relies on the ability to 'construct and manipulate external representational means, such as diagrams or equations' (ibid: 695). Recent work by Marion Vorms (2011) and Axel Gelfert (2016) has developed this

[15] An inferentialist view of scientific models has been defended by Swoyer (1991), Hughes (1997), Suárez (2004), and Contessa (2007). The epistemic tools approach has been proposed by Boon and Knuuttila (2008), Feest (2010), and Knuuttila (2011).
[16] See also Knuuttila (2011).

point, highlighting Feynman diagrams as a good illustration of how 'the use of external representational means – even when their material dimension consists only in marks on paper, combined with conventions regarding their use and interpretation – facilitates abstract reasoning and stablilizes inferences by anchoring them in a more accessible medium' (Gelfert 2016: 113-114). On the other hand, although abstraction is often included as one of the main tools in constructing models, there is a temptation to reify this feature into a more substantive claim about the ontology of models vis-à-vis the claim that models *are* abstract entities existing in some Platonic realm. However, the pragmatist perspective discussed here provides no grounds for thinking that models are, in themselves, abstract objects (Kuorikoski and Lehtinen 2009: 123).

The core idea of an inferentialist approach to model-based understanding is to understand models in terms of their inferential properties and how they enable and license certain kinds of inferences. What does this mean? On the one hand, it exemplifies the functional approach to models introduced earlier (e.g., Keller 2000); that is, the focus here is on how models are used and function in our reasoning practices. The inferentialist approach is less concerned with what models *are,* and more with what they are *used for*. Rather than analysing some real-target system, scientists often construct models which allow them to circumvent the complexity inherent to a real-world target or phenomenon. These model systems are constructed from explicit and implicit conceptual commitments, sets of inferential rules, and a range of inferential connections between claims about the target system and claims about the source system. The activity of modelling amounts to reasoning through the implications and consequences of the assumptions and commitments that are built into the model. As Johannes Lenhard argues they allow scientists to explore the implications of adopting specific theoretical assumptions (Lenhard 2009: 178).

From the inferentialist perspective, the epistemic value of a model is to be found in its capacity to sustain, extend and articulate our inferential reasoning with regards to the phenomena of interest and to enhance our inferential ability. On this view, models are analysed first and foremost as tools that enable scientists to draw inferences that may not for various reasons have been otherwise possible with human reasoning powers alone. In some cases this is because of the complexity of the phenomena being investigated. In others, the issue is the scale of the computational power required. Models thus function to reduce the cognitive workload of the model-users, by simplifying the object of analysis and providing external aids for our thinking. This idea has been linked to recent work in cognitive science, which identifies the role of *cognitive scaffolding* through the use of various strategies and external tools to support our inferential abilities (Clark 2011; Toon 2015). Doing so enables scientists to narrow 'the space of information search by localizing the relevant features of the object in a perceptually salient and manipulable form' (Boon and Knuuttila 2011: 315). De Donato

Rodríguez and Zamora Bonilla explain this idea as follows, where idealized assumptions fill the role of 'virtual commitments':

> […] the complex set of factual claims and inferential dispositions constituting the body of commitments of a scientific community has the *potential* of generating a lot of consequences, but the particular connections between many of those claims and dispositions may make it very difficult, or even impossible, to *actually* derive these consequences by using the calculating and reasoning techniques the community has […] In these cases, the introduction of a definite set of "virtual commitments", depicting a counterfactual situation (that perhaps is even incompatible with some accepted principles), may act as a "canal system", or a "pruning" that channels the inferential flow from our inferential dispositions toward more interesting and numerous logical consequences. (de Donato Rodríguez and Zamora Bonilla 2009: 108)

This illustrates something we have already encountered in Brandom and Rouse. Scientific communities, disciplinary fields or research groups operate on the basis of a number of conceptual commitments about what the world is like and how to understand phenomena within their particular domains. As Brandom argues, the open-endedness of our concepts typically commits us to far more than we know how to say, and the complexity of the range and scope of the inferential consequences of these commitments may be heightened in the sciences. The introduction of a model, which is constructed on the basis of these conceptual commitments, provides a tool for teasing out these inferential consequences, for determining which inferential consequences are possible candidates for endorsement, and for working out what our concepts actually commit us to. On this reading, scientific models can be thought of as tools that enable scientists to ask what the inferential consequences of our conceptual commitments *would be* if they were understood in this sense rather than another. Hence, modelling depends upon our ability to reason counterfactually through our concepts, to try and work out what they would say about the world *if* they were true.

Furthermore, Kuorikoski and Ylikoski (2015) argue that the external character of models and their expression in various external representational means helps to increase the *reliability* of our inferences. On the one hand, both psychology and everyday experience show that unaided human reasoning is not very reliable. We are victims to limitations in our working memory, various biases, and errors in our reasoning. By externalising our inferences, we increase their reliability (2015: 3824). On the other hand, they argue that the acts of writing down equations or coding computer programs force scientists to be explicit about the assumptions they are making about the target and their conceptual commitments. Doing so makes the validity of the inferences drawn easier to check.

Linking the construction and analysis of models to our inferential abilities underpins the idea that models provide understanding through their role in conceptual articulation. It is

because models have these features that conceptual articulation – working out the appropriate circumstances in which our concepts apply and determining their inferential consequences – is possible. Furthermore, the epistemic tools view means that conceptual articulation is not an ethereal disembodied matter of reasoning, but embedded in particular concrete formats that structure how scientists make sense of and understand the world.

5.2 Intelligibility

Models facilitate understanding by rendering their target phenomena intelligible and conceptually meaningful. The notion of intelligibility has been a prominent topic in recent debates on understanding. However, it has been primarily characterised as a subjective quality that scientists project onto a body of knowledge or theory. Riggs (2003), for example, argues that while intelligibility is connected to understanding, it only captures our ability to discern order and connections between things, but requires no substantial connection to truth. Grimm (2011) furthers this point by holding that intelligibility only captures 'subjective understanding', as merely a grasp of how specific propositions interlink but one that does not depend on their truth or factivity. De Regt (2009a, 2015) proposes a more robust conception of intelligibility, but restricts it to 'the value that scientists attribute to the cluster of qualities of a theory that facilitate its use' (de Regt 2015: 3793). As such, intelligibility is primarily conceived of as a value that scientists project onto theories.

These treatments of the notion of intelligibility within the literature on understanding are problematic for two reasons. First, intelligibility has been framed as a property of theories or models that facilitate their use in understanding. It is thus treated as a feature of the vehicles of understanding (the theory or model) rather than of the objects of understanding (the target or phenomenon). However, understanding a phenomenon and understanding a model should be distinguished. Models are not just used because they are intelligible to scientists, but because they help to make the phenomena themselves intelligible. Second, intelligibility is not a contingent feature of our understanding of things. As Heidegger and Brandom argued, intelligibility is an essential condition on understanding anything at all. If something is unintelligible, I cannot claim to understand it.

This point can be illustrated with respect to the development of classical genetics in the first decades of the twentieth century, focusing on the work of T.H. Morgan and his colleagues on the model organism *Drosophila melanogaster*. Rouse (2009, 2011, 2015a) argues that the significance of this period in the development of genetics lies in the fact that it is with the modelling work of Morgan that 'genes' shifted from being merely hypothetical posits to being the locus of a whole field of inquiry, namely genetics (Rouse 2009: 45). The conception of genes as the principal objects of study marks out this period as a distinct conceptual

development in our understanding of genetics. Although the idea of a 'unit' of heredity had formed in the earlier work of Mendel, Weismann, Bateson and others, in nineteenth-century studies of heredity, genetic phenomena could not be distinguished 'from the larger processes of organismic development within which they functioned' (Rouse 2015a: 303).[17] Importantly, this conceptual development was not exclusively a theoretical matter, but was underpinned by the development of new practices, such as the production of pure lines and cross-breeding (Rheinberger 2010: 154). According to Rouse, by correlating meiotic cross-over frequencies of mutant traits with visible transformations in chromosomal cytology, the *Drosophila* group enabled significant and extensive inferential articulation of the concept of a gene, and in doing so opened up the field as a novel research domain (Rouse 2015a: 304).[18]

As we saw in Chapter Two, inferentialist approaches to conceptual meaning and intelligibility hold that concepts acquire meaning through their use in judgements and in their relation to other concepts. This view of conceptual meaning and understanding is thus holistic: conceptual content is expressed inferentially by being connected to other concepts through various material, counterfactual and incompatibility relations. Articulating our conceptual understanding involves working out the inferential relations in which they are connected to one other, as well as determining their depth and scope. The use of *Drosophila* as a model organism illustrates how conceptual articulation takes place in scientific practice. By being able to connect specific phenotypic traits with chromosomally located 'genes', the use of *Drosophila* enabled extensive conceptual articulation. To illustrate this, Rouse offers the example of the judgement in classical *Drosophila* genetics that the Sepia gene is *not* on chromosome 4:

> This judgement does not simply withhold assent to a specific claim; it has the further content that either the Sepia gene has some other chromosomal locus, or that Sepia mutants vary in more (or less) than one 'gene'. Such judgements, that is, indicate a more or less definite space of alternatives. Yet part of the content of the 'simpler' claim that Sepia is on chromosome 3 is the consequence that it is not on chromosome 4. (ibid: 304)

The point that Rouse takes from this is that such a judgement is intelligible only because it presupposes the intelligibility of an entire conceptual space which connects traits, loci, and genes, 'including the boundaries that mark out what is or is not a relevant constituent of that space' (ibid: 304). Intelligibility, from this perspective, is not simply a value that scientists

[17] For the theoretical background to Morgan's work, see Bowler (2003). For the development of *Drosophila* as a model organism, particularly with respect to its connection to the molecular gene concept, see Weber (2005: ch. 6).

[18] Rouse's and Rheinberger's reading of classical genetics can also be seen to challenge a dominant theoretical interpretation of classical genetics as being first and foremost about a theory of transmission. For discussion, see Waters (2004).

assign to particular theories or models, which then enables them to use them more effectively. Instead, it operates at a more basic level and concerns the way in which the phenomena being studied can be brought into a scientific 'space of reasons', which 'encompasses not only the claims that we take to be true or false but also the conceptual field and patterns of reasoning within which those claims become intelligible possibilities whose epistemic status can be assessed' (Rouse 2015a: 17). The use of *Drosophila* as a model organism allowed scientists to articulate the concept 'gene' in a way that enabled distinctly genetic phenomena to become intelligible, so that scientists could investigate, discuss, and reason about them. Only then could claims about genetic phenomena become truth-apt (Hacking 2002). Before we can ask whether our concepts and the theories and models that employ them adequately represent the world, or whether and to what extent they are justified or true, we must already have some grasp of what they say about the world.

5.3 Modality

A second significant feature of the use of models in conceptual articulation concerns the way in which they afford modal understanding of their target phenomena. This modal understanding refers to an understanding of the possible ways in which phenomena *could* be. As we saw in chapter two, Sellars (1948) argued that empirical concepts are lawlike insofar as their inferential scope and our grasp of them spans over counterfactual possibilities: 'If this thing I have in my hand *were* a piece of copper, then it *would* conduct electricity'. For Sellars, our understanding of empirical concepts is counterfactually robust. Working out the counterfactual and modal possibilities inherent to the phenomena being modelled constitutes an important aspect of model-based conceptual articulation.

A useful example here is the Schelling checkerboard model from economics and the social sciences. Originally developed by Thomas Schelling in *Micromotives and Macrobehaviour* (1978), the model is an attempt to understand and explain why cities are often ethnically segregated with different ethnic groups living in different parts of a city.[19] It is widely held that such segregation is the result of either strong discriminatory preferences (e.g., racism) or else the result of economic factors such as welfare differences among different groups (Aydinonat 2007: 429). The checkerboard model focuses on the effect of discriminatory preferences on the neighbourhood structure to understand how a segregated neighbourhood might result from these preferences.

[19] The Schelling checkerboard model is a stock example in the philosophy of social science literature and has been discussed by several authors. Here, I draw on the discussion of this model as described in Aydinonat (2007), Ylikoski and Aydinonat (2014), and Kuorikoski and Ylikoski (2015.)

Schelling's idea was to assume that there are two types of agents, As and Bs, which can be taken as individuals or families, and who populate the checkerboard-like neighbourhood. In this neighbourhood, some slots/squares are empty and available for occupation. The agents are assumed 'to discriminate between their own group and other groups' (Ylikoski and Aydinonat 2014: 20). They are, moreover, assumed to have certain preferences concerning the ethnic composition of their neighbourhood: As may prefer to live together with Bs, while Bs may prefer to live only in neighbourhoods without As. A distinction is also assumed between 'strong' and 'mild' preferences. The former is for the agent to wish to live in a neighbourhood where its own type is the majority. The latter, mild preference, implies that the agent is happy to live in a neighbourhood as a minority. However, it is supposed that agents will be unhappy if they have extreme minority status, which is where their own type is less than one-third of the neighbourhood (ibid: 20). The idea is that agents will move to new (randomly assigned) slots until they are satisfied with the composition of their neighbourhood. Doing so changes the composition of the neighbourhood the agent leaves and that of the neighbourhood they join. In other words, if an A moves out of a neighbourhood, the number of As in that neighbourhood decreases, and the number of As in the new neighbourhood increases. The process ends when every agent is content with its neighbours (ibid: 21). Schelling found that ultimately agents end up in neighbourhoods consisting of all the same type. The primary inference that can be drawn from this is that even mild discriminatory preferences can lead to strongly segregated neighbourhoods in the efforts of agents to avoid being in a clear or extreme minority.

Emrah Aydinonat has argued that the conceptual significance of the Schelling checkerboard model is that it highlights some of the mechanisms and factors that *may* bring about residential segregation and some of the possible ways in which they may interact (Aydinonat 2007: 430). What the model shows is that segregation could occur even in cases where racist attitudes were not assigned to the inhabitants of a neighbourhood. This point is developed further by Kuorikoski and Ylikoski (2015), who argue that the Schelling checkerboard facilitates scientific understanding by enabling scientists to make correct 'what-if-things-had-been-different' inferences about the world. They argue that because of its capacity to sustain counterfactual what-if inferences, the Schelling checkerboard model has explanatory value.[20]

The modal nature of inferences afforded by models highlights the way in which understanding with models can articulate an expanded space of possibilities in terms of which

[20] Kuorikoski and Ylikoski argue that the checkerboard model provides understanding because it has this explanatory import. While I agree with their conclusion, my account of model-based understanding and conceptual articulation differs since I argue that understanding is already provided if the model articulates intelligible inferential possibilities.

the target phenomenon is understood. This is important because it suggests how actual understanding may nevertheless depend on an understanding of what is possible, a point that was anticipated in Heidegger's account of understanding as a 'projection onto possibilities', but which more recently has been defended by Peter Lipton (Lipton 2009: 50). Although one might object that scientists should be seeking characterisations of phenomena in actual terms, this would be too restrictive a desideratum when we consider the nature of scientific inquiry. The 'how-possibly' understanding afforded by modelling can form a crucial part of the research process. As Carl Craver and Lindley Darden argue in the case of mechanism schemas in the life sciences, determining what is possible for a given phenomenon is often heuristically useful in directing research programmes in the design of experiments:

> In such cases, the scientist has to make a decision about whether the idealizing assumptions matter (that is, make a difference) for the purpose to which the schema is to be put. A literally false schema, that is, might be plausible enough for a given purpose even if it is not, technically, a how-actually scheme. Scientists often stop short of developing how-actually schemas and are satisfied instead with identifying a region of the space of possible mechanisms within which the actual mechanism falls. Idealizing models often help to identify that region of the space of possibilities, the space of possibilities that are how-actually enough. (Craver and Darden 2013: 35)

This is not to claim that how-possibly models and explanations are necessarily to be favoured over models which provide how-actually explanations. Instead, the point here is that the development of more refined, how-actually models and accounts of phenomena often crucially depends on the establishment of an intelligible space of possibilities within which the research process can be focused. The counterfactual inferential possibilities afforded by models function to constrain and delimit the range of factors and mechanisms that must be taken into account to arrive at an understanding of a phenomenon. Furthermore, Craver and Darden suggest that between how-possibly models and how-actually models lies a range of *how-plausibly* models, which are developed as the number and variety of constraints on our understanding of the target phenomenon increases, thereby gradually rendering some how-possibly models implausible.

Recognising that modal understanding can be provided by models through counterfactual inferences also points to the possibility of an epistemically productive role for 'failed' models. Nancy Cartwright and Sindhuja Bhakthavatsalam (2017: 452) suggest this with the case of the Rutherford model of the atom. This model represents the atom as made up of a central charge, the nucleus, surrounded by a cloud of orbiting electrons. From both a theoretical and empirical perspective, the model is now known to be grossly inaccurate. According to modern quantum mechanical models, the electron does not revolve around the nucleus like planets around the sun in a solar system, as the Rutherford model suggests. Furthermore, it is also empirically

inadequate insofar as it predicts that the electron will continuously lose energy and spiral into the nucleus causing the atom to collapse. Nevertheless, the Rutherford model still affords some degree of understanding. Although we have now rejected the model, Rutherford's model initially represented a greater degree of understanding than previous models (it was an advance, for example, on Thomson's plum pudding model). However, more importantly, it shows how an atom *could not* be; that is, it illustrates a *physical impossibility*. Thus, on the one hand, the model dictates how some inferences are *incompatible* with a correct understanding of the atom, for example, the inference that the electron orbits around an atom's nucleus like a planet around the sun. On the other hand, it suggests a level of counterfactual understanding: '…*if* electrons revolved around the nucleus the way the model says they do, then matter *couldn't* exist as we know it' (ibid: 458). Thus, even if some models are theoretically or empirically inadequate, this does not necessarily mean that they no longer provide understanding. That may suggest why models that are known to be false are still used in science education: not only do they contain some truth with respect to older models, but they also illustrate the kinds of inferences that are *prohibited* with regards to specific concepts (ibid: 459). The key here is how models articulate what it is possible to legitimately *say* about such phenomena. Thus, the modal understanding afforded by models has normative import: they prescribe how we *ought* to think about the phenomena and constrain the range of legitimate possibilities that are available to us.

The modal understanding afforded by models is perhaps most vividly illustrated by considering the construction of models of phenomena that are known not to exist. Weisberg, for example, cites the work of R.A. Fisher, who suggested that 'the only way to understand why there are always two sexes involved in sexual reproduction is to construct a model of a three-sexed sexually reproducing population of organisms' (Weisberg 2007: 223; citing Fisher 1930). Only by constructing a model of such a phenomenon is it possible to study such counterfactual possibilities: 'Modelers are often interested in phenomena such as three-sex biology, perpetual motion machines, or non-aromatic cyclohexatriene because, insofar as we can understand why these phenomena do not exist, we will have gained a better of [sic] understanding of phenomena that do exist' (ibid: 223).

Nevertheless, focusing on the modal or counterfactual nature of the understanding afforded by models raises two objections. The first complaint would be that understanding modal possibilities only has heuristic value at best, whereas what is required is an understanding of how things *actually* work.[21] Does a determination of what is possible qualify as

[21] This is implicit, for instance, in Craver and Darden's schema of how-possibly, how-plausibly, and how-actually models. The ultimate goal in mechanism research is arriving at a how-actually model.

understanding? Second, one might object that a focus on the modal understanding afforded by scientific understanding risks being too permissive. It might suggest that the ability to infer any counterfactual consequence that follows from the model would qualify as understanding. The issue can be seen as a descendent of Nelson Goodman's (1954) 'grue' problem: since many inferences and inferential rules are compatible with the same body of data, what justifies projecting a concept one way rather than another? However, I think these objections miss the wider picture.

I think these objections miss the wider picture: with regards to the first objection, I have already indicated that understanding of modal possibilities has a central and legitimate role in scientific research, and that it fundamentally enhances our understanding of target phenomena. As such, it should be treated as a cognitive success. Schelling's model for example was concerned to show that segregation is not necessarily the result of racist attitudes. The model suggests that it is possible for segregation to arise even if people would prefer to live in non-segregated neighbourhoods. More substantially, the modal understanding afforded by models also licenses the development and construction of other models to examine and explore the possibilities presented by the original model. Again, the reason that the Schelling model had conceptual significance was because its implications for our understanding of socioeconomic phenomena extended beyond its own application. Various iterations of the Schelling model and other related models systematically varied and modified the assumptions (concerning, for instance, variation in the measure of preferences regarding the composition of one's neighbourhood) built into the model. This enabled economists and social scientists to expand the scope of the kind of possible influences and effects that would result in the segregation equilibrium.[22] In this respect, the initial modal understanding provided by the Schelling model motivated and functioned as a guide for the direction of future research.

Regarding the second objection, models are not created out of nowhere; they are constructed and analysed within specific disciplines that are guided by certain norms of intelligibility and background understanding that delimit the space of legitimate counterfactual inferences. Cartwright and Bhakthavatsalam argue that the reason why Schelling's checkerboard model provides counterfactual understanding is because the conjecture that segregation can result from even mild discriminatory preferences is at least plausible based on our familiarity with real people and their preferences, and thus can be treated as consistent with the real world in key respects (Cartwright and Bhakthavatsalam 2017: 458). In this sense, the counterfactual space of possibilities engendered by the model is constrained by its degree of familiarity, plausibility and intelligibility, and because of this, it is reasonable to assume that not all *logically* possible counterfactual inferences would or

[22] See Kuorikoski and Ylikoski (2015: 3823) for further references.

should be endorsed. In this respect, the kinds of constraints on the inferences one can draw from models echo Kuhn's analysis of thought experiments: models work by drawing on familiar uses of concepts and intelligible possibilities in the world. Furthermore, the possible inferences that one could derive from the construction and analysis of models are constrained by the connection of those inferences to other inferential commitments that are already endorsed by a scientific community. Not just any counterfactual inference will be accepted by the community; rather, the kinds of counterfactual possibilities that a model suggests should be thought of as candidates for inferential commitments that the community would be prepared to endorse. Those inferential consequences will, moreover, be evaluated by the extent to which they cohere with our current conceptual commitments, their facilitation of future research, and correspondence with further experimental results and observations. In that respect, modal understanding is not just a case of 'anything goes'.

6 Conceptual Normativity and Model-Based Understanding

The final aspect of model-based conceptual articulation that I wish to discuss concerns its normative dimension. Models are constructed on the basis of certain inferential norms; norms which connect the model with its target, and norms which prescribe the inferences one can draw within the model (e.g., Suarez 2015). While model-based inferential norms can be made explicit by scientists and codified in publications or in educational contexts, most norms are implicit in scientific practice (de Donato-Rodriguez and Zamora-Bonilla 2009). These norms often exist in the form of ready-made methods, formats or techniques for analysing phenomena, such as simulation techniques (Winsberg 2006), mathematical methods (Knuuttila and Loettgers 2015) or computational templates (Humphreys 2004). The Schelling model, for instance, provided a natural template for further manipulation and investigation of agent-based accounts of segregation. Similarly, the Lotka-Volterra model,

> […] resulted from an attempt to apply to a new field some modelling methods and techniques taken from (for them) exemplary and more fundamental disciplines […] [T]he available mathematical representational tools and the way they were used in mechanics guided [Volterra] in imagining and describing the predator-prey system in a particular way. Thus some already-established mathematical tools and modelling methods functioned as scaffolding for Volterra's (and Lotka's) scientific imagination. (Knuuttila and Loettgers 2015: 22-23)

In this respect, models can be thought of as loci of conceptual commitments and inferential norms. Their application and extension in the form of established methods, techniques and templates suggests they can be constitutive of the norms governing scientific inquiry and practices across different disciplines (de Rodríguez and Zamora Bonilla 2009: 106). However,

the simple recognition of such inferential norms and their place in modelling practice does not capture what is at issue when considering the question of conceptual normativity. The issue, rather, is how such norms and rules are established in the first place and come to be endorsed within a community, as well as their connection to scientific understanding. Note that the issue here is *not* that of determining the appropriate epistemic norms of model-based reasoning, in the sense of determining how the beliefs or claims derived from models are rationally warranted. Again, to repeat the familiar theme, whether or not something can be assessed for its warrant or truth depends on it already being manifest as an intelligible possibility. The assessment of the truth, justification or factivity of the detected pattern depends upon whether or not the pattern is considered as a *candidate* which can then be open to assessment. The issue therefore concerns the normativity of conceptual meaning: what it takes for a model to say something meaning about a phenomenon in the first place. The normativity arises from the possibility of a model not showing or saying anything meaningful about its target phenomenon at all.

To illustrate this point, it is helpful to return to the analogy between experiments and models (Mäki 2005; Morgan 2003). As we saw above, modelling and experimentation are both attempts at *isolating* the casually relevant factors. In the case of experimentation, such isolation is achieved causally through experimental controls. In the case of models, such isolation is achieved using idealized assumptions. In this sense, experimentation and modelling are both attempts to introduce relatively well-behaved circumstances in order to discern phenomena. However, 'phenomena' should not be thought of as just something 'Given', in the sense of Sellars' (1997) 'Myth of the Given'. Instead, the notion is akin to Bogen and Woodward's (1988) conception of phenomena as stable regularities in the world, or Hacking's conception of a phenomenon as something 'salient and noteworthy' in experimental contexts (Hacking 1983: 220).[23] For both conceptions, phenomena are distinguishable from what is observed or observable in particular cases. Both have the character of an orderly and regular pattern against a background of complexity, where this pattern holds across different contexts. The challenge of empirical and experimental work, Hacking emphasises, is being able to set up an experiment in the right way in order for phenomena to manifest themselves as salient or noteworthy. Building on this, Rouse argues that the import of conceiving of phenomena in this way concerns the significance of what such discernible patterns tell us about the world:

> Phenomena *show* something important about the world rather than our merely *finding* something there […] Creating a phenomenon is an achievement. The focus of that

[23] For further discussion of the relation between these two conceptions of phenomena and their relation to experimental practice, see Feest (2011).

> achievement…is the salience and clarity of a pattern against a background […] Some natural events have the requisite salience and clarity, but Hacking argued that most do not. Other phenomena must therefore be created in laboratories or other sites of scientific intervention. (Rouse 2015a: 226)

I suggest that models can be thought of as 'sites of scientific intervention'. Theoretical model systems are ways in which scientists can introduce relatively well-ordered circumstances (through the use of idealized assumptions) in order to create conditions in which intelligible patterns in the phenomena can manifest themselves. For example, the Lotka-Volterra model system of predator-prey interaction could show that predators will thrive when the prey population is high, but that eventually the predator population will outstrip the number of prey and go into decline. This was on the basis of idealized assumptions, such as the intrinsic exponential growth rate for the prey in the absence of predators and a constant death rate for the predators, combined with established methods and techniques from mathematics and physics. This pattern is not something that is just given; rather it is precisely through the construction and analysis of a model that such patterns can be detectable in the first place.

From this perspective, the normativity of modelling pertains to the extent to which a model can reveal not just any pattern in the world, but *significant* patterns. As Rouse argues, what makes phenomena conceptually significant is 'that the pattern they embody is informative beyond its own occurrence' (ibid: 233). This can be illustrated with respect to the work of the Morgan group on *Drosophila* in classical genetics:

> Their experiments correlated differences in crossover frequencies of mutant traits with visible differences in chromosomal cytology. If these correlations were peculiar to *Drosophila melanogaster*, or worse, to these particular flies, they would have had no scientific significance. Their salience instead expressed the sense of a more general pattern in the cross-generational transmission of traits and the chromosomal location of 'genes' as discrete causal factors. (ibid: 233)

Rouse's point is that, at this stage, the relevant normative measure is not whether we are justified in accepting that a pattern discerned by the construction of a model is actually there, but whether and in what respect that pattern is worth exploring as a defeasible projection of stable invariance across different and unexamined cases. In this respect, the significance of modelling practice is its ability to discern intelligible and significant patterns in target phenomena that motivate and guide ongoing and future research. To reiterate: this is not a matter of our psychological or cognitive capacity to actually draw out the inferential consequences derived from a model; instead, the issue is a normative one. It concerns which inferential consequences – which discernible, significant patterns – we *commit* ourselves to and take *as* inferentially significant. The relevant sense of commitment here is normative in the Brandomian sense: in committing to a pattern detected by model analysis, we make a claim

that the pattern initially picked out by the model is one that *ought* to hold in other, unexamined cases. In committing ourselves to the significance and intelligibility of a pattern, we take on a responsibility for its rational defeasibility; that is, *saying* that a significant pattern is there at all means placing that phenomenon and what the model says about it in the scientific space of reasons, opening it up to rational assessment, and taking upon ourselves the justificatory responsibility to offer reasons for it. This in turn feeds into Rouse's normative conception of practices. For in the case of modelling practice, what is *at issue* is how to explore and make sense of the manifest pattern, and what is *at stake* is whether that pattern is meaningful at all. That there is something at stake in the significance and intelligibility of a pattern is the possibility of that pattern being merely apparent or coincidental.

These considerations concerning the capacity of models to conceptually articulate phenomena connect with scientific understanding in at least one relevant sense through scientific skills and capacities for pattern recognition. Here, Rouse draws on Haugeland's analysis of pattern recognition as an essential skill for understanding a domain of entities. The critical point is that pattern recognition is normative, rather than a fact about our psychological capacity for discernment:

> To speak of recognition is to allow for the possibility of error […] [T]he patterns that show up in phenomena must not merely indicate a psychological or cultural propensity for responsiveness to them. Our responsiveness to them, our *taking* them as significant, must be open to assessment. (ibid: 232)

Haugeland's analysis of pattern recognition identifies two related but fundamentally different sorts of pattern recognition:

> On the one hand, there is recognizing an integral present pattern from the outside – *outer recognition* […] On the other hand, there is recognizing a global pattern from the inside, by recognizing whether what is present, the current element, fits the pattern – which would, by contrast, be *inner recognition*. The first is telling whether something (a pattern) is *there*; the second is telling whether what's there *belongs* (to a pattern). (Haugeland 1998: 285)

For Haugeland, outer recognition concerns our ability to tell whether a pattern is there: to understand that the relevant dependency relations between aspects of a modelled phenomenon constitute a possible pattern. The relevant sense of scientific understanding in the case of outer recognition is being able to recognise a pattern as a possible candidate for pattern recognition and our grasp of the relation of that pattern to a broader space of intelligible possibilities. This is why the capacity of models to afford modal understanding is crucial: outer recognition depends on our understanding of what is possible for phenomena within a particular scientific domain. Inner recognition is the ability to tell whether particular instances count as an element

in the larger pattern; to be able to tell whether a particular case is in accord or at odds with the larger projected pattern. In Haugeland's example: 'When chess players recognize pieces, positions, and moves, it is outer recognition of the constituted phenomena; when they recognize these phenomena as legal or illegal, it is inner recognition of the constitutive standards' (Haugeland 1998: 285-286). In other words, playing chess requires on the one hand an understanding of what the legitimate possibilities for chess phenomena, such as pieces and moves, are. What it means for something to be a Queen, for example, is for it to correspond to a certain pattern of behaviour (in accord with the rules of chess); there is a requisite pattern of behaviour that constitutes something functioning as a Queen in chess. On the other hand, players must have the ability to discern when a particular instance of using the Queen is in accord with that projected pattern, and to respond appropriately if a move does or does not conform to the projected pattern.

The reason Haugeland thinks there must be *two* kinds of pattern recognition is to allow for the possibility that the two come into conflict, because it is only in the discrepancy between the two that the possibility of the projected pattern *being accountable to something beyond itself* emerges. In other words, the crucial point is that the possibility of discrepancy between the two allows for the possibility of our claims about the world being *objective*: 'What is crucial for objectivity is that the two recognitive skills be distinct […] In particular, skilful practitioners must be able to find them in conflict – that is, simultaneously to outer-recognize some phenomenon as present (actual) and inner recognize it as not allowed (impossible)' (ibid: 286). It is only because what our theories and models say about the world can come into conflict with how the world actually is that they can be accountable to the world. In the face of such conflict, our theories, models, scientific skills, and practices then become subject for repair, revision or rejection (Haugeland 1998: 336-339). It is only because there is a gap between how we *take* the world to be and how the world actually *is* that our models say anything meaningful about the world in the first place.

This discussion of pattern recognition connects models to scientific understanding in two ways. First, by articulating the world conceptually: the construction and analysis of models enables the phenomena they discern to become intelligible. The relevant sense of intelligibility (or meaningfulness and significance) is determined by the extent to which a model can detect a projectible pattern of behaviour. Such patterns are thus presented as intelligible candidates for endorsement, in the sense of committing to that pattern as something that *ought* to hold in other, unexamined cases. The conceptual normativity of models pertains to the significance of this pattern and whether or not it is worth exploring it as a possible candidate for saying something about how the world is. As such, models can say something about the world by bringing phenomena into the scientific space of reasons, and only then can the claims derivable from the model become open to assessment.

Model-based scientific understanding therefore concerns, on the one hand, the possibility of whether what the model says about its target phenomena is meaningful or intelligible in the first place. On the other hand, the understanding provided by the model concerns not only our ability to recognise that pattern of behaviour in unexamined cases. It also concerns whether particular cases are in accord with the pattern projected by the model, or whether it is fundamentally at odds with it. Particular cases that violate the pattern lead to the possibility of revising what exactly that pattern says about the world, whereas continued recalcitrance presents the possibility of questioning whether there was ever a meaningful pattern in the first place. The projected pattern sets up a norm, as it were, and our understanding involves being able to tell whether particular cases are in accord with that norm or violate it. It is only because our understanding involves this grasp of what it would take for something to violate what our models say about the world that our model-based understanding can be accountable to the world.

7 Concluding Remarks

We have been considering the ways in which models play a role in the process of conceptual articulation. First, models are often a central feature of rendering a domain of phenomena intelligible. I argued that the intelligibility provided by models should not be restricted to a value that scientists project onto them. Instead, the kind of intelligibility that models provide pertains to the phenomena themselves. For example, the standardisation of the model organism *Drosophila melanogaster*, combined with experimental practices such as chromosomal mapping, provided a way of opening up and conceptually articulating the field of genetic phenomena. The relevant sense of intelligibility here refers to the use of models to bring such phenomena into the 'scientific space of reasons'. It is only because models can bring phenomena into the space of reasons, that we can then be in a position to judge whether what models say about those phenomena can be assessed in terms of truth or falsity, or whether the inferences they make possible are justified.

Second, I argued that models, such as the Schelling checkerboard model of segregation, also afford modal understanding of their phenomena, insofar as they are often constructed for the purpose of delimiting a space of intelligible possibilities for making sense of specific phenomena. Prior to determining whether the inferences derived from models are true, we must already have a grasp of what they say. Therefore, by focusing primarily on the contribution of models to the constitution of conceptual meaning and significance, models afford a more basic level of understanding than other current accounts of model-based

understanding, which have so far been restricted to questions concerning the truth and factivity of idealized models, have acknowledged.

Third, I focused on the normative dimension of model-based understanding. Drawing on Rouse and Haugeland, I argued that the normativity of modelling emerges in at least two respects. The first concerns the possible significance of the patterns discerned in phenomena. The issue here is in what sense the detected pattern holds and whether it is informative beyond its own domain of application. The possibility of a pattern holding beyond a model's own domain of application can be thought of as staking a claim about how the world ought to be, and as such it becomes a candidate for assessment. It is at this stage that models can be tested in more familiar terms: one can test their predictive or explanatory power, for instance, their coherence with our standing conceptual commitments, or the possibility of yielding further inferential consequences that can be questioned and assessed. The second aspect of normativity concerns our understanding more directly and shows up in our ability to tell whether particular cases of the pattern projected by a model count as instances of that larger pattern. Crucially, scientific understanding requires being able to tell when the two conflict.

The significance of this discussion is that it prompts reassessment of the ways in which models facilitate scientific understanding. I have argued that models play a crucial role in articulating the intelligibility of scientific phenomena. Furthermore, this role has a fundamental normative dimension: models articulate possible inferential patterns that a scientific community ought to be endorsed. As such, the understanding provided by models pertains not simply to the way in which models represent how the world is. Instead, it involves recognising how unexamined cases projected by the model are in accord with or violate the norm established by that inferential pattern. Therefore, my contention is that an exclusive focus on the factivity or non-factivity of model-based understanding is too narrow, for it fails to acknowledge the way in which models provide understanding through their constitution of conceptual meaning, prior to their assessment of whether what they say is warranted or true-or-false.

Chapter Five

Rethinking Explanatory Understanding

1 Introduction

Within the philosophy of science, much of the recent interest in scientific understanding has been focused on the connection between explanation and understanding. This reflects a widely held belief that the ideal of scientific understanding is explanatory understanding, since explanations provide us with a correct account of why something is the case. Yet there is little agreement within the current literature on how to characterise explanatory understanding. Kelp (2015) draws a helpful distinction between two broad camps. On the one hand, *explanationists* argue that it is knowledge or evaluation of explanations alone that provide scientific understanding. On the other hand, *manipulationists* think that explanatory knowledge by itself is not enough; instead, understanding requires the ability (or abilities) to manipulate certain representations or concepts embedded within an explanation.

This chapter addresses an assumption that is shared by both sides of the debate. The assumption is that explanatory understanding is appropriately analysed at the level of the *individual subject*. I will argue that this individualistic approach to explanatory understanding is inadequate for characterising the connection between explanation and understanding in scientific practice for two reasons. First, scientific understanding is not just about what happens in individual minds but is constituted by the distributed and collective abilities of scientists to reason about and investigate objects within research communities across disciplinary contexts. As such, I propose to address the *social* dimension of explanatory understanding. Second, explanationists have argued that the skills or abilities proposed by manipulationists are either trivial or unnecessary for explanatory understanding. I challenge this claim by providing an alternative conception of the skills required for understanding, one that builds upon and extends the analysis of the social dimension of understanding.

I proceed as follows. Section Two expands upon the claim that the current debates have assumed an individualistic stance towards explanatory understanding. In Section Three, I build upon Andrea Woody's (2015) recent work on the role of explanation in scientific

practice, which suggests that the relation between explanation and understanding is social and cognitive in complex, interdependent ways. For Woody, explanations do not simply pick out certain dependencies between phenomena. Instead, they perform crucial conceptual work in constituting and perpetuating the norms of intelligibility for a scientific domain. Section Four extends this idea by appealing to John Haugeland's analysis of the normativity of scientific understanding. Haugeland's analysis suggests that the skills required for understanding turn on our responsiveness to the constitutive norms and standards of a domain of entities. As such, I argue that this provides a robust conception of the role of skills in scientific understanding that does not reduce them to individual psychological abilities. Building on Woody and Haugeland, I argue that explanations help to shape the normative landscape in which scientific understanding is situated. I thus locate the connection between understanding and explanation at the level of norm-governed, social practices.

2 Understanding and Explanation

In the philosophy of science, it has long been taken for granted that the gold standard of scientific understanding is explanatory understanding. Explanatory understanding is understanding *why* something is the case. We understand why, for example, light is refracted as it enters water once we have the right kind of explanatory details regarding that phenomenon. Traditionally, philosophers of science have tended to assume that possessing an explanation leads automatically to understanding. As Nounou and Psillos point out:

> The Standard Story goes like this: Scientific understanding is constitutively tied to explanation; hence, it is covered by theories of scientific explanation. Bluntly put, the question is this: what kind of information should science offer (and how should it offer it) in order for it to provide understanding of the world? And the standard answer is: it should provide explanatory information. (Nounou and Psillos 2012: 72)

Nounou and Psillos claim that explanations provide understanding in virtue of their providing *explanatory information*; in other words, explanatory *content*. The standard models of explanation – the deductive-nomological model (Hempel 1965), the causal-mechanist model (Salmon 1984), the unificationist model (Friedman 1974; Kitcher 1989) – all offer competing accounts of what this explanatory content is. However, no matter what the precise nature of the explanatory information and the binding relation between a given *explanans* and its corresponding *explanandum* turns out to be, it is generally agreed that once these details are worked out, we will *ipso facto* acquire an account of explanatory understanding. These basic assumptions form the basis of what Kelp (2015) terms the *explanationist* approach to scientific understanding.

The explanationist approach has been articulated, developed and defended in recent years by Kareem Khalifa (2012, 2013) and Michael Strevens (2013). Building on the general picture provided by the Standard Story, Khalifa, for instance, identifies explanatory understanding with knowing that (a) the *explanans* and (b) the *explanandum* are true, and (c) what the correct explanatory link is between *explanans* and *explanandum* (Khalifa 2012: 26). On this view, explanatory understanding does not involve anything over and above the knowledge of the explanatory information contained within the *explanans* and the *explanans-explanandum* relation. Explanations are the vehicles for explanatory understanding, and it is therefore assumed that we can defer to the literature on explanation and the various models of explanation that have been proposed over the years in order to make sense of scientific understanding.

Recently, the explanationist approach and its defence of the Standard Story has been challenged from two directions. First, it has been argued that knowledge of an explanation is too strong a condition for understanding. Lipton, for example, reviews a number of cases where it is plausibly argued that we can acquire scientific understanding without an explanation. To establish this claim, Lipton distinguishes between explanations on the one hand and the cognitive benefits of explanations on the other. He then identifies understanding a phenomenon with having the cognitive benefits of an explanation of it and distinguishes between four types of cognitive benefit an explanation may offer: knowledge of (i) causal information, (ii) necessity, (iii) possibility, and (iv) unification (Lipton 2009: 43-44). In order to show that one can have understanding without explanations, he argues that one can come to enjoy these benefits of explanation without at the same time having an explanation of the relevant phenomenon. If this is the case – if one can understand a phenomenon without having an explanation - then explanationist accounts of understanding are too demanding. The intellectual demands they place on understanding – knowledge of an explanation – are unrealistically high. As such, Lipton argues that less sophisticated or rigidly formulated cognitive achievements can qualify as understanding.

The second challenge to the explanationist camp comes from the *manipulationists*. The manipulationist's basic claim is that explanatory understanding requires more than knowledge of an explanation; rather, explanatory understanding essentially involves some kind of ability (or abilities) to manipulate certain representations. The motivation for this view stems from the idea that one might have all the relevant explanatory information for a phenomenon, while still not understanding why that phenomenon is the case. For example, asked 'Why do planes fly?', a student might know that Bernoulli's principle and the relevant features of the plane, while still not understanding why planes fly. First, she might know Bernoulli's principle but never have attempted to apply it to the question of why planes fly. Or, second, she might lack the ability to do so (Toon 2015: 3862; de Regt 2009a: 26). As Kuhn and others have pointed

out, students are often able to recite important theoretical principles, and yet have no idea how to apply them or how such an application might work in practice (Kuhn 1970: 23-51).

The manipulationists hold that what the student is missing is specific kinds of *skills, abilities* or *know-how* that need to accompany the relevant explanatory knowledge. For de Regt (2009a), it is the possession of specific epistemic skills that enables an individual to extract the right information from an explanation in order to acquire understanding: 'The extra ingredient needed to construct the explanation is a *skill*: the ability to construct deductive arguments from the available knowledge' (de Regt 2009a: 26). Similarly, Grimm (2010: 340 - 341) argues that understanding-why requires that a subject S be able to make predictions for new variations on previously studied examples.[1] A variation of this appeals to the idea of grasping: understanding requires that we also 'see' or 'grasp' how facts, principles and explanatory knowledge fit together. Thus, Catherine Elgin writes, 'to understand the Comanches' dominance of the southern plains involves more than knowing the various truths that belong to a comprehensive, coherent account of the matter. The understander must also grasp how the various truths relate to each other' (Elgin 2009: 323).

It should be noted that the differences between the explanationist and the manipulationist camps are not always sharp. Indeed, we find authors in either camp incorporating or blending elements of the other into their accounts. For example, those who fall into the manipulationist camp, like de Regt (2009a) or Grimm (2011), do not deny that knowledge of explanations is involved in scientific understanding, and in fact, they are in significant agreement with the basic thrust of the Standard Story. For example, de Regt's criterion for understanding phenomena still involves explanation as an essential condition (de Regt 2009a: 32). But they do insist that some extra cognitive abilities are required before the possession of that knowledge qualifies as understanding why something is the case. Similarly, those who defend the explanationist view, such as Khalifa (2012, 2013), recognise that some abilities are required. He argues that the ability to evaluate or distinguish between good explanations is certainly required, but that *pace* the manipulationists, this does not require a new theory of understanding, but is already captured within the literature on scientific explanation.

As the debate stands, it is unclear how the dispute between explanationists and manipulationists could be resolved. There is no agreed upon characterisation of explanatory understanding. Ultimately, the worry is that the disagreement simply reflects the fact that different writers appeal to different intuitions about what explanatory understanding is. Manipulationists address the sceptical worry that one might possess all the relevant

[1] Grimm adapts this from Woodward's (2003) manipulationist account of causal explanation, according to which the measure of understanding is the ability to answer 'What-if-things-were-different?' questions. A similar view is defended in the epistemological literature by Alison Hills (2015).

information and knowledge concerning facts, phenomena, and theoretical principles, but still not understand why something is the case. Explanationists, by contrast hold that all that needs to be said about explanatory scientific understanding is adequately captured in the literature on the epistemology of scientific explanation. From the explanationist perspective, the worry is that it is not at all clear why we need a new account of understanding in the first place. Khalifa, for example, argues that the manipulationist thesis is redundant, and that its putative insights have already been expressed in the literature on scientific explanation. He cites with approval the following passage from Hempel, which supposedly undercuts the intuitions of the manipulationists. Hempel says: 'The understanding [that scientific explanation] conveys lies in *the insight* that the explanandum *fits into or can be subsumed under*, a system of uniformities represented by empirical laws or theoretical principles' (Hempel 1965: 337; quoted in Khalifa 2012: 26; italics are Khalifa's). Khalifa's idea here is that the subject of understanding does not require some special ability, skill, or know-how. Such skills and abilities, he argues, are already implicit in the literature on the epistemology of scientific explanation (Khalifa 2012: 27). The propositional knowledge contained in a correct explanation is enough to generate understanding.

The debate risks descending into an exchange of examples and counterexamples designed to show that the other side has either missed or over-emphasised certain features of explanatory understanding, in a way akin to how the literature on scientific explanation has been plagued by philosophers quarrelling over a set of reputed, but still disputed, counterexamples: the flagpole and the shadow, the ink spill on the carpet, and the hexed salt, to name but a few. In the scientific explanation literature, the worry is that the dispute between explanationists and manipulationists cannot possibly be settled in this manner because the counterexamples to putative claims to explanation rest on the same core intuitions that created the disparities between viewpoints in the first place. If you have already accepted that causal information is the key to explanatory understanding, you are unlikely to be moved by counterexamples that suggest that causal information is not always sufficient. It is not difficult to see that a similar situation risks hampering the debates over the nature of explanatory understanding: a cynical reader might worry that rather than making any serious progress on the elucidation of explanatory understanding, theorists of understanding can expect to bat around competing intuitions concerning students, Bernoulli's principle, and Comanche history for some time. The likely outcome would be a stalemate.

But beyond these methodological challenges, I want to suggest that a deeper assumption shapes the debate in equally unhelpful ways. Common to both sides of the debate is the presumption that the adequate focus of analysis is at the level of the *individual subject*. This explains why understanding has traditionally not been treated with much interest by philosophers of science. Precisely because understanding was only ever thought of as a

subjective or psychological state that is idiosyncratic to individuals, it was regarded as irrelevant to making sense of scientific explanation (Hempel 1965). Nevertheless, even while rejecting the narrow psychologistic reading of the concept of understanding, this assumption has filtered through into a recent edited volume on scientific understanding (de Regt et al. 2009). Although they recognise the social nature of scientific understanding, when considering the subject of understanding, de Regt et al. (2009: 3) write that 'this subject is typically a scientist who understands a phenomenon'.

This individualism informs both camps in the debate. For example, in Khalifa's explanationism, understanding is first and foremost a mental state possessed by the individual epistemic subject. It is this recognition that prompts Khalifa's shift of the discussion away from explanation itself and towards the epistemological component of explanations. Thus, the relevant mental state is knowledge of an explanation, which amounts to 'rich, accurate, and detailed beliefs about an explanation' (Khalifa 2012: 18). Similarly, the manipulationists primarily treat the kinds of abilities relevant to understanding as psychological abilities pertaining to the individual (e.g., Grimm 2011). This is perhaps most vividly seen in Daniel Wilkenfeld's (2013) account of understanding as representation manipulability. Wilkenfeld proposes that a condition on understanding is the mental representation requirement: 'in order to understand some object *x*, a thinker must possess a mental representation of *x*' (Wilkenfeld 2013: 1003). This requirement is then incorporated into the following manipulationist account of understanding:

> URM (Understanding as Representation Manipulability): A statement, attributed in context C, that thinker T understands object o, is true if and only if T possesses a mental representation R of o that T could (in counterfactuals salient in C) modify in small ways to produce R', where R' is a representation of o and possession of R' enables efficacious (according to standards relevant in C) inferences pertaining to, or manipulations, of o. (ibid: 1003-1004)

The suggestion then is that both camps have taken it for granted that an adequate characterisation of explanatory understanding should proceed on individualistic terms. Now, of course, focusing on how understanding is achieved by the individual may not in itself be a complaint if our aim is to make sense of human understanding in general, and my point is not to dismiss that approach altogether. That said, the conception of understanding that I have developed in this thesis suggests that it is not possible to focus on characterising the individual subject of understanding without reference to its place within social practices.

Furthermore, if we seek an adequate account of *scientific* understanding, a focus solely on the individual is unhelpfully limited, particularly if we consider how scientific understanding is manifest in modern scientific practices. At least as far as a modern conception of the sciences goes, scientific research and inquiry is not just the result of individuals working

together. At the social level, as recent work in social epistemology suggests, knowledge exhibits different features than at the individual level. Questions of authority and expertise, collaboration, the distribution of warrant and accountability, and shared epistemic values suggest quite a different picture (Goldman & Blanchard 2015). This would suggest that *scientific* understanding is embodied in communities of practitioners in highly complex and interconnected ways (Galison 1997). In addition, the advent of 'big science', in which large groups of scientists coordinate their labour to conduct research, carried out through multidisciplinary projects occurring within international networks, suggests that an account of the social dimension of explanatory understanding would be more reflective of contemporary scientific practice (e.g., Shrum et al. 2007).

The emergence of such extended and globalised research communities has significant philosophical and epistemological implications for the status of understanding at this socially distributed level. Similarly, building on work in social epistemology, questions that we might ask about group knowledge might also be relevant to the idea of group understanding (Bird 2010; Leonelli 2014; Wray 2001). This point is also anticipated in recent work in philosophy of science that draws on ideas concerning extended and distributed cognition in cognitive science (Giere 2006; Nersessian 2006; Ylikoski 2009), as well as in recent scholarship on collaborative research (Wray 2002; Andersen 2016; Ankeny & Leonelli 2016). Therefore, if we seek to provide an adequate account of explanatory understanding, specifically in the case of how it shows up in scientific contexts, then we will need to pay attention to the essentially social nature of scientific understanding. My proposal therefore is to locate explanatory understanding at the level of social practices.

A related assumption that is rarely questioned in the current debates on explanatory understanding pertains to how explanations themselves are conceptualised. This assumption is that to determine what characterises explanatory understanding, or what kinds of features it has, our primary aim should be to make sense of the content of the *explanans* and the nature of the relation that binds it to its corresponding *explanandum*. That is, we take as our starting point, the standard formal-logical or ontic conceptions of scientific explanation. For example, for de Regt et al. '…while explanation may be viewed as a two-term relation between an *explanans* (for example, a theory) and an *explanandum* (the phenomenon), understanding is always a three-term relation involving *explanans*, *explanandum*, and a subject' (de Regt et al. 2009: 3). This assumption overlooks the broader roles of explanatory discourse in scientific practice.

Accounts of explanatory understanding typically characterise understanding as a kind of propositional attitude towards the content of an individual explanation. Accordingly, the debate between the two camps in the recent debates on understanding turns on the question of whether we need to appeal to new features or aspects that have not been traditionally been

given as much attention in the epistemology of scientific explanation (skills, know-how, abilities, counterfactual reasoning) to make sense of that propositional attitude, or whether the epistemology of scientific explanation already captures everything we need to know about understanding. This point is illustrated by Khalifa's Explanatory Model of Understanding, according to which, 'Any philosophically relevant ideas about scientific understanding can be captured by philosophical ideas about the epistemology of scientific explanation without loss' (Khalifa 2012: 17).

Traditionally, the literature on scientific explanation has predominantly (if not exclusively) focused on answering the question: 'What is *the* nature of scientific explanation?' Furthermore, it is typically assumed that the appropriate answer will reveal the common logical and/or conceptual structure of individual scientific explanations. We can see this in existing accounts of explanation, which may be broadly characterised to fall under two categories (Salmon 1984). *Epistemic* accounts ascribe an essential role to the vehicles by which explanations are conveyed, and typically take explanations to be arguments. The two best-known examples are Hempel's deductive-nomological account and the unificationist account. In the deductive-nomological account, explaining something (e.g., why light bends in water) consists in showing how the *explanandum* follows deductively from a law (laws of refraction) and background conditions (density of the medium). Hempel's approach to scientific explanation provides us with a *prescriptive* account of explanation, in the sense that something only counts as explanatory if it meets certain formal requirements. In unificationist accounts, as proposed by Friedman and Kitcher, explaining something consists in showing how an *explanandum* which falls under certain laws can be subsumed under laws of wider scope.

By contrast, ontic accounts of explanation focus on showing how the *explanandum* fits into natural patterns or regularities. The focus here is for the explanation to pick out a causal process or mechanism that makes a difference to the *explanandum*. On these accounts, a scientific explanation is an objective account of how the real world is connected by showing how causes and mechanisms produce their effects. An explanation is only warranted if it discloses the causal structure behind the given phenomenon. The ontic view stems from (Salmon 1984) and is currently most popular in the guise of mechanistic accounts (Machamer et al., 2000; Craver 2007; Craver and Darden 2013), or Woodward's interventionism (Woodward 2003).

Philosophical accounts of scientific explanation have thus predominantly focused on the binding relation between *explanans* and *explanandum* to characterise the nature of scientific explanation. However, this kind of approach has always been troubled by the diversity and plurality of things that count as explanatory within scientific practice. Within recent studies of scientific practice, analysing the material investigative practices of scientists and tracking

the distinctive and heterogeneous patterns of explanatory reasoning found in various sub-disciplines of the science has motivated forms of scientific pluralism (Kellert, Longino & Waters 2006; Chang 2012). Explanation is not just one kind of thing; there exist many different types of accounts which scientists regard as explanatory. In the natural sciences, for example, there exist not only nomic accounts, but also mathematical, probabilistic, causal, functional, and structural accounts. In addition, philosophers of science have also turned to the social sciences and the humanities, in which we find intentional and interpretive explanations. These provide information about either motives or meanings relating to agents and their interactions within social groups. Recognising the plurality of different types of accounts that are treated as explanatory has thus led recent studies of scientific explanation to reject the idea that explanations are a homogenous kind.

Similarly, pragmatic theories of explanation have argued that explanations cannot be divorced from the pragmatic contexts in which they occur and emphasise the relevance of the explainer's interests and background knowledge. Pragmatic accounts treat explanations first and foremost as appropriate answers to explanation-seeking questions and hold that it is the pragmatic context that determines not only what would count as an appropriate answer, but also helps to fix what is relevant within a particular explanation (Faye 2007, 2014). From this perspective, explanations are not just products that can be assessed in terms of the formal-logical or objective structure, but should be seen in the context in which they motivate developing or refining an agent's belief system. The extent to which explanations are successful depends on their coherence with an agent's background knowledge, and in relation to the epistemic goals of discursive inquiry. The pragmatic view does therefore not try to prescribe universal conditions for what makes something explanatory; it instead recognises that scientific explanation is 'goal-oriented and context-bound' (Faye 2007: 44).

The upshot of these remarks is that there is good reason to think that both sides of the debate concerning explanatory understanding need to take seriously the connection between explanation and understanding as it is manifest in ongoing scientific practice. As such, I have identified two main issues in the debates concerning explanatory understanding. The first concerns the predominant focus on the individual subject of understanding. The second concerns the extent to which skills and cognitive abilities are an essential condition for explanatory understanding. To address these problems, I shall argue that turning to the role of explanation and explanatory discourse in scientific practice acknowledges the fundamentally social, pragmatic and normative dimension of explanatory understanding.

The core of my argument is that explanatory understanding cannot be divorced from our norm-governed, social practices. To motivate this, I begin with Andrea Woody's recent functional approach to scientific explanation in Section Three. Woody's approach argues that explanations, and explanatory discourse more generally, fulfil an important normative

function in scientific inquiry by constituting the norms and standards of intelligibility of scientific communities and disciplines. In Section Four, I extend Woody's analysis by turning to Haugeland's account of the constitutive normativity of scientific understanding. Both accounts, I shall argue, help to show that the connection between explanation and understanding operates in relation to these constitutive standards and norms of intelligibility: understanding is the ability to recognise how something accords with those standards and norms, and explaining something is showing how it fits in with those standards and norms.

3 Rethinking Explanation: Woody's Functional Perspective

3.1 The Functional Perspective

Woody begins by identifying three questions that we might ask of explanation:

1. What are the adequacy conditions for individual scientific explanations?
2. How should explanatory power be justified as a theoretical virtue, if indeed it should be?
3. What role(s) does explanation play in science?

Woody observes that an answer any of these questions puts constraints on answers to the other two. She contends that philosophers have focused much of their attention on question (1) and have only then proceeded to (2) and (3). As we saw in the previous section on the standard accounts of scientific explanation in the literature, the prevailing focus on the *explanans-explanandum* has been driven by the implicit assumption that this is the primary focus of concern: that statements should be accepted as explanatory if, and only if, the right sort of relationship exists between the *explanans* and the *explanandum*. Furthermore, with regards to (2), the justification of explanatory power as a theoretical virtue has so far mainly been approached through debates concerning the legitimacy of inference to the best explanation. The third question is rarely considered.

Woody argues that we reorient our discussions of scientific explanation around question (3). Drawing on Hempel's work on functional analysis, which is an attempt 'to understand a behaviour pattern or a socio-cultural institution by determining the role it plays in keeping the given system in proper working order' (Hempel 1965). Woody proposes that we conceive of explanations as discursive activities which are situated within scientific communities. The emphasis is placed on explanation as 'an activity, or custom, of scientific communities, rather than the propositional content produced as a result' (ibid: 80). Woody thus proposes the

functional perspective on scientific explanation as an attempt '[…] to reveal how the practice of explanatory discourse functions within scientific communities given their more comprehensive sets of aims and practices' (ibid: 80). In comparison with traditional accounts of scientific explanation, explanation is thus treated first and foremost as a social activity with particular social functions.[2] Consequently, the functional perspective asks us to 'think about where and when explanations are sought and formulated' and to 'consider what role(s) they might play in practice' (ibid: 81).

Woody offers five generalisations that frame the functional perspective:

> 1) Explanatory discourse (in contrast to explanations in the abstract) is social in nature, in the sense that it is an activity between at least two agents (the explainer and the explainee).
> (2) Explanatory discourse and reasoning is mediated through differential levels of authority and expertise, such that the legitimacy of an explanation may depend more on the authority of the explainer than the content of the explanation.
> (3) Scientific disciplines and sub-disciplines typically take for granted sets of exemplary explanations. These operate as templates that guide inquiry and expectations concerning phenomena within a specific domain.
> (4) Explanatory discourse frequently involves articulating communally accepted explanations, which often take the form of repetition of known explanations, rather than the cultivation of novel ones;
> (5) Explanations play a central role in scientific training and education.

Woody contends that from this list we can begin to get a sense of the kinds of situations and contexts in which explanations are sought and formulated. What is key here is that a functional analysis of the kinds of exemplary explanations that underpin different scientific disciplines serve to make salient the *norms of intelligibility* governing those communities and disciplines, which can be tracked historically in the evolution and transformation of research communities. For example, explanatory discourse often involves the communication of exemplary explanations, which serve to guide what a scientific community judges to be explanatory, and what models of intelligibility that community has chosen to embrace (Woody 2003: 24). Similarly, within educational contexts, one aim of explanatory practice is training novice

[2] It might seem that a focus on explanation as an activity was already heralded by the pragmatic theories of explanation advocated by Bas van Fraassen (1980) and Peter Achinstein (1983). While there certainly is much overlap between the two approaches, the classic pragmatic accounts of explanation are still predominantly concerned with the adequacy conditions for what makes something count as an explanation, and thus as far as the functional perspective is concerned, still operate within the confines of the traditional project as set out above.

practitioners to recognise typical explanatory patterns. Scientific textbooks, as well as being repositories of explanatory information and content, function as instruction manuals which help to communicate norms and explanatory standards to new members of a community (ibid: 23).

In this sense, Woody argues that the production of explanations serves to *constitute* rather than merely *communicate* intelligibility for a scientific discipline. Thus,

> […] because explanatory discourse inculcates particular patterns of reasoning, it functions to sculpt and subsequently perpetuate communal norms of intelligibility. In effect, explanations encode the aims and values of particular scientific communities, telling practitioners what they should want to know about the world and how they should reason to get there. (ibid: 81)

From this perspective, explanations, and explanatory discourse more broadly, are not exhausted by the specific kind of content or explanatory information embedded within particular explanations. Rather, explanations have a crucial *socially prescriptive* role through coordinating scientific disciplines and research communities on the basis of specific aims, values and norms of intelligibility. This socially prescriptive role helps to explain the persistence of the convenient fictions (e.g., frictionless planes), exemplars (e.g., the Punnett square analyses of classical genetics), and idealizations used within explanations across the evolution of different research communities and disciplines, despite their known falsity or the existence of more accurate refinements. Woody cites with approval Cartwright's (1983) discussion of why certain *ceteris paribus* laws in the physical sciences, like Snell's law, are kept on the books and still taught:

> [T]he reasons have to do with the task of explaining, specifying which factors are explanatorily relevant to which others is a job done by science over and above the job of laying out the laws of nature […] we have to *decide* what kinds of factors can be cited in explanations. One thing that *ceteris paribus* laws do is to express our explanatory *commitments*. They tell what kinds of explanations are permitted. (Cartwright 1983: 47-48; emphasis added by Woody)

The suggestion therefore is that the presence of such fictions, exemplars and idealizations fulfils an important social and normative function in constituting and perpetuating the norms of intelligibility that underpin the practices of research communities and their disciplinary identity. This in turn suggests how such elements help to coordinate scientific research by guiding judgements of explanatory relevance. Woody contends that judgements of explanatory relevance are rarely determined by individual scientists acting alone, but rather depend upon their training and assimilation within particular social practices and communities. As such, explanatory commitments 'are not, at their foundation, commitments of individuals' but rather 'commitments of communities' (Woody 2015: 81).

The case for the functional approach to scientific explanation is enhanced by a continued increase in large teams of international scientists that coordinate diverse types of data and modes of analysis. The existence of such globalised institutional structures requires scientists to be able to communicate reliably across wide geographic, social, and cultural distances. To maintain the coherence of scientific activity and research under these conditions means that the 'actions, aims, and assumptions of communities of practitioners must be robustly coordinated and aligned' (ibid: 82). Explanatory discourse is one tool for meeting this challenge. As Alan Love argues, 'Explanatory discourse is the scaffolding that holds this unruly mass together and channels it productively' (Love 2015: 91).

Woody's suggestion is *not* that philosophical analysis of the epistemic or ontic nature of scientific explanations should be abandoned. Her suggestion is that we start from scientific practices, and build up from there to the formal and logical analysis of explanations.[3] In turn, her account suggests that while manipulationists were right to argue that explanatory understanding involves more than the possession of the right kind of explanatory knowledge, they have misidentified what this extra feature is. The function of explanatory discourse is to inculcate particular patterns of reasoning as exemplary, thereby sculpting and perpetuating communal norms of intelligibility. By doing so, explanatory discourse enables coherent practice across different practitioners, communities and sub-communities. The value of the functional perspective lies in its potential to reveal how explanatory discourse supports the particular epistemic and practical aims of given scientific communities. To illustrate these points, we can now turn to one of Woody's main examples, the Ideal Gas Law.

3.2 The Ideal Gas Law

The functional account of scientific explanation highlights several aspects of explanations and explanatory discourse that are missed or overlooked by traditional conceptions of explanation, for example, the role of explanatory discourse in articulating and communicating shared standards and norms of intelligibility for scientific disciplines. To demonstrate how these features of explanatory discourse operate, Woody focuses on the Ideal Gas Law in chemistry.[4] The Ideal Gas Law is particularly useful in the context of investigating the relation between

[3] In that respect, the philosophical task for explanation would be more akin to a form of 'explanatory expressivism' in Brandom's or Sellars' sense, in which philosophical accounts of scientific explanation make explicit and codify the explanatory norms that remain implicit in scientific practice.

[4] Elsewhere, Woody has applied a functional approach to the periodic table (Woody 2014) and explanations from molecular geometry using diagrammatic orbital schemes (Woody 2004).

explanation and understanding, as it has been used as an example in several discussions of scientific understanding.[5]

The Ideal Gas Law is based on the following gas laws, which are considered special cases of the general law:[6]

> *Boyle's Law*: The volume (V) of a given mass of gas at a constant temperature (T) is inversely proportional to its pressure (p). $pV = constant_B$
> *Charles' Law*: The volume (V) of a given mass of gas at a constant pressure is directly proportional to the absolute temperate (T). $V = constant_C\ T$
> *Avogadro's Law*: At a given temperature and pressure, equal volumes of gas contain equal numbers of moles. $V = constant_A\ n$

For a given mass of gas kept at a constant volume, the pressure law states that the pressure is directly proportional to the thermodynamic temperature. These gas laws are combined in the Ideal Gas Law: $pV = nRT$, where n is the number of moles, and R is the universal molar gas constant. Thus, on its surface the Ideal Gas Law is an equation of state relating three variables: pressure, volume, and temperature.

What is an ideal gas? An ideal gas is one that obeys the gas laws. It is a hypothetical gas, that is, one in which all collisions between atoms or molecules are perfectly elastic and in which there are no intermolecular attractive forces. Thus, ideal gas molecules occupy negligible space and have negligible forces between them. To derive the Ideal Gas Law, scientists generally make the following idealising assumptions (Mizrahi 2012: 242):

> (GL1) Gas molecules are moving in random directions in the sample. If, for some reason, the gas molecules were all moving strictly up and down (or from left to right) in the container, then the law would not hold.
> (GL2) Interactions between gas molecules can be neglected. In some gases, there are unusually large attractions or repulsions between molecules. For the most part, however, the average distance between molecules in a typical gas sample is so large that these interactions have negligible effects.
> (GL3) The average energy of gas molecules is proportional to the temperature of the gas.

Thus, in its modern form – that is, combined with its interpretation on the basis of idealising assumptions - the Ideal Gas Law is a paradigmatic *ceteris paribus* law (Cartwright 1983),

[5] For example, Elgin (2007), de Regt (2009a), and Mizrahi (2012).
[6] Drawn from Mizrahi (2012: 241-242).

which means that interpreted literally, it applies only to ideal gases – gases in the limit of zero pressure – which of course do not exist. *Prima facie*, this raises a dilemma for the law: interpreted literally, the law is either false or applies to no particular gases depending on whether we explicitly articulate the *ceteris paribus* clause. But if that is the case, how we do account for the law's significance in chemical and explanatory practice?

One option is to treat the law as a convenient approximation which is used liberally to model a vast array of relationships among properties of actual gases; that is, to use it as an approximate description of the relations between physical magnitudes (Woody 2004: 19). In that respect, the Ideal Gas Law is treated as a phenomenological approximation with decent empirical adequacy.[7] However, Woody argues that restricting the law's value in chemical practice to providing roughly accurate empirical descriptions of the quantitative relations that obtain between actual gas properties fails to capture its significance and explanatory status. That is, the legitimacy of its function within scientific practice does not rest on the (phenomenological) accuracy of its approximations.

Instead, Woody argues, the conceptual role of the Ideal Gas Law is not merely descriptive, but also prescriptive: it coordinates reasoning about gases by focusing chemists' attention on specific properties under a conception of how to think about gases as 'particulate in nature, being composed of small, extremely dense particles that exert negligible forces on one another by virtue of large average separation distances' (Woody 2013: 1574). Woody's point is that the explanatory and conceptual function of the Ideal Gas Law in chemistry extends beyond its approximate accuracy in describing gas phenomena. The law is used in chemical explanations because the idealization involved in the law effectively selects a set of properties ($pV = nRT$) as conceptually significant for chemical practice and understanding:

> The law [expresses] a conception of the core theoretical properties of gases. The relevance of an ideal gas model for understanding actual gas behaviour would not depend directly, then, on the empirical accuracy of the approximations included in the ideal gas description. Not merely an inaccurate description, the law provides selective attention to certain gas properties and their relations by ignoring other aspects of actual gas phenomena. It instructs chemists how to think about gases *as gases*. (Woody 2004: 21)[8]

By idealising from the original empirical generalisations expressed in Boyle's, Charles' and Avogadro's laws, the Ideal Gas Law yields a widely applicable mathematical framework that generates a standard for recognising deviations from the 'ideal' expressed by the equation, which in turn become *explananda* for chemistry: 'By trying to account for the failure of gases

[7] Note that 'phenomenological' here is intended in the scientific sense, where a phenomenological model describes the empirical relationship between phenomena.

[8] In that sense, the Ideal Gas Law provides, in Heideggerian terms, an articulation of something akin to the intelligibility (being) of gas entities *as* gas entities.

to obey Boyle's law exactly, we can learn about the size of molecules and the forces that they exert on one another' (Mahan 1975: 33; cited in Woody 2004: 24). In this respect, the idealization 'guides model construction for actual gases, and when the law poorly approximates actual gas behaviour, properties are conceptualized as deviations from the ideal' (Woody 2015: 82). Therefore, the Ideal Gas Law in chemical practice is not simply descriptive of dependencies between phenomena, but prescriptive: it instructs chemists how to conceptualise and understand gas behaviour, provides 'selective attention to, and simplified treatment of, certain gas properties', and articulates patterns of explanatory reasoning about gases that extend beyond its initial application.

Additionally, the Ideal Gas Law facilitates two-way inferential traffic between macroscopic and molecular properties. Because the Ideal Gas Law is derived from the kinetic theory of gases, the idealized representations of gas properties follow from idealized conceptions of gas particles. The 'zero pressure' limit of an ideal gas links directly to the theoretical assumption that gas particles exert no forces on each other, while at the same time, because the assumption that zero pressure is achieved by large distances between particles, it suggests the system can be modelled as a collection of Newtonian point masses undergoing elastic collisions only with the container (Woody 2015: 82). What this suggests is that the Ideal Gas Law 'serves as a bridge between the realm of bulk substances, the traditional subject matter of chemistry, and the realm of atoms and molecules, the discipline's endorsed theoretical framework' (Woody 2004: 24).

This in turn, Woody argues, provides the foundation for the concept of 'temperature'. The idea here is that because of our subjective and inherently comparative perception of heat, raw experience cannot provide an adequate grounding for a coherent, quantitative concept of temperature. By contrast, because the Ideal Gas Law stipulates a linear relationship between volume and temperature, it articulates a partial definition of what temperature is: namely, a property that under constant pressure varies directly, and precisely, with gas volume. As a result, the Ideal Gas Law imposes structure on empirical domains, rather than passively describing them (ibid: 23; Woody 2015: 82).

The Ideal Gas Law and its function within explanatory discourse and reasoning thus helps to illustrate why a functional perspective reveals a more complex, cognitive role for explanation and explanatory discourse within scientific practice, extending it beyond traditional conceptions and analyses: rather than just being a particular kind of *explanans*, the Ideal Gas Law performs crucial conceptual and normative work within chemical practice. It 'provides scaffolding for reasoning about gas behaviour, it facilitates the theoretical unification of macro and micro level descriptions, and grounds our understanding of core concepts such as temperature' (ibid: 82). The central role of this idealization as a scaffold for explanatory reasoning in chemistry helps to account for its place in opening chapters of

general textbooks, providing a model of intelligibility that not only provides propositional content and information regarding the behaviour of gases, but also orients the entire field of chemistry, both conceptually and methodologically (ibid: 82) by connecting empirical observations with its underlying theoretical framework.[9] The functional perspective asks us to pay attention to actual scientific practices and holds that doing so foregrounds the 'socially normative status of explanations…Explanations are revealed to be a primary means of inculcating shared exemplars that function as norms of intelligibility' ibid: 84).[10] As such, Woody concludes, explanations invoking the Ideal Gas Law remain common in chemistry and scientific practice precisely because the Ideal Gas Law prescribes an ideal of intelligibility for the entire discipline.

3.3 The Functional Perspective: Virtues and Challenges

Both in the way in which it frames the debate (the focus on functional roles of explanation in scientific practice rather than adequacy conditions of explanation as such) and the terms by which explanation is analysed, the functional perspective represents a significant departure from traditional conceptions of explanation, and, I shall argue in the next section, from current conceptions of explanatory understanding. Yet, because of this, there are worries one might have about this approach.

Firstly, one might argue that the functional perspective leaves us with no criteria for judging the adequacy of explanatory practice. Given Woody's rejection of the search for universal criteria and the emphasis on the social normative function of explanation, one might worry that it leaves us with a mentality of 'anything goes', or that it portrays scientific practice as no more than a pernicious form of Kuhnian 'groupthink'. If all that matters for certain explanations is that they fit patterns of reasoning endorsed by the community, is normative evaluation possible? Woody's response to this is that normative evaluations are available, but not in the fashion to which philosophers of science might be accustomed. The functional perspective permits normative evaluation of explanatory adequacy, but this is now framed in terms of comparing patterns of explanatory discourse with explanatory aims. That is, the locus of evaluation is the fit between such patterns and the discipline's goals. Adequacy of proposed explanations is judged on the basis of 'whether a given pattern or model of intelligibility is

[9] For example, in a recent introductory chemistry textbook, the first section is devoted entirely to the Ideal Gas Law and its bearing on our understanding of gas phenomena. See Atkins and de Paula (2014).

[10] Woody explicitly connects explanations with exemplars in the Kuhnian sense: they enable normal science to proceed and establish conditions for membership and accountability within disciplinary communities.

well suited to support the cognitive and epistemic aims of the respective scientific community' (Woody 2015: 83). Mechanistic explanations, for example, are ill-suited to both the theoretical foundation and predictive orientation of quantum-theoretic scientific practice.[11] Such normative judgements will thus be contextual and more geared towards engaging with and benefiting ongoing research.

Second, one might reply that this response does not really address the issue. While it provides a sense in which explanations can be judged in terms of their fit with a community's epistemic aims, this does not seem to provide an effective way of distinguishing between *genuine* explanations and those merely accepted as such by the community. The functional perspective could respond that such a request for a way of distinguishing between genuine and false explanations lies outside its remit; its orientation, that is, is towards making sense of what functions as explanatory discourse in scientific practice as it finds it. A stronger response would, however, be that if we seek an account of explanation as such, we need to take into consideration how and why our best scientific practices treat some things as explanatory and others as not. We should not presume that the only way to make sense of explanation is to characterise a set of objective, two-place relations that determines which facts in the world explain other facts, without risk of begging the question. The point then is that these determinations between genuine explanations and false ones must be answered through the practices that generate exemplary explanatory reasoning, which seem to yield significant resources for judging the adequacy of competing explanations. As Woody notes,

> To label something an explanation is not to identify it as a particular sort of descriptive information, but rather to endorse it, to suggest that it provides the right way to think about things, relative to certain (often implicit) aims. Well-established explanatory patterns hold such honorific status and function as exemplars of reasoning that bind practitioners together in ways that cultivate and support the social coherence and productivity of particular scientific endeavors. (ibid: 84)

Of course, this should not mean that simply because a community endorses an explanation that it therefore is automatically justified or true. The problem with this is that, if this were the case, the scientific community could not be mistaken in adopting particular explanatory standards. As I argued in the previous chapter, there has to be a distinction between how we take things to be and how things actually are. Our norms of intelligibility and the explanations that make sense in the light of them must be accountable to something beyond themselves in

[11] As Peter Dear (2006) argues, the response of practitioners to such explanations turns less on their lack of empirical adequacy or theoretical justification, but more on the fact that explanations do not even make sense in such a context. Similar conceptual disputes about what counts as explanatory are evident throughout early modern scientific practice, as argued in Gaukroger (2006, 2012).

order to not be conceptually empty or vacuous. Woody's point, I take it, is that treating something as explanatory is to suppose that it offers good, but defeasible reasons for thinking about phenomena in a certain way. This suggests how explanations might be accommodated within a broader inferentialist framework. From the inferentialist perspective, the act of providing an explanation would be interpreted as providing *reasons* for why others within a community should take up particular commitments with regards to the *explananda*. On this model, an explanation would be akin to providing (possibly deeper) reasons for a certain conclusion which is need of justification. This does not mean that it cannot be challenged, revised or refuted; instead, in endorsing a particular explanation, we take responsibility for it in the 'game of giving and asking for reasons' and propose it as a candidate for endorsement within the community, which can then be subject to empirical testing, scrutiny and evaluation. In that respect, an inferentialist view would, like Woody's functional perspective, emphasise the normative pragmatics of scientific explanation.

It is the fact that explanations can take on a normative status within a scientific community that illustrates the social nature of explanatory understanding. Proposed explanations offer putative entitlement to a scientific community to infer certain conclusions. Julian Reiss (2012), for example, has proposed an inferentialist analysis of causal claims. According to Reiss, causal claims, such as 'Smoking causes lung cancer' can be understood as a way of licensing certain kinds of inferences to other related claims, such as 'Smoking raises the probability of lung cancer' and 'If John hadn't smoked, his chance of developing lung cancer would have been considerably smaller' (Reiss 2012: 770). In turn, Reiss argues that such claims can be challenged or scrutinised; whether a scientific community is entitled to make an inference depends 'on what the background beliefs and the purposes and goals of the investigation are, and on what is known about the evidence and the causal relation in question (ibid: 772). The broader significance of this is that it highlights the thoroughly social nature of explanatory discourse:

> […] entitlement [to an explanation] is social because there is no guarantee that an individual holds all the beliefs that are jointly necessary to warrant an inference […] Causal claims in the social and, more ostensibly, the biomedical sciences are often established by what has come to be known as 'big science': large numbers of researchers who bring different kinds of expertise to a common research project […] [N]o individual researcher has all the evidence required to justify knowledge claims. Rather, numerous scientists are required, each one with his or her specific expertise. (ibid: 773)

The suggestion then is that explanations do not just convey information. Accepted explanations take on a normative significance within scientific practice. Thus, Woody claims several virtues for a functional reorientation in philosophical analyses of scientific explanation. First, by rejecting any kind of essentialism regarding the adequacy conditions for

scientific explanation, the pragmatic orientation of the functional perspective accommodates the variety of apparent explanations observed in contemporary scientific practice. Second, despite the responses to the above challenges, it preserves insights from the philosophical literature on scientific explanation. However, rather than treating them as competitors for the one true, universal account of scientific explanation, it recasts them as localised and context-specific explanatory norms that operate in conjunction with the various epistemic aims of research communities. Third, the functional perspective offers a reappraisal of the notion of explanatory power:

> The diversity of our examples makes it implausible that we might conceive of explanatory power as a straightforward property of a theory or model […] In each of our examples the explanatory structures do not supply the most theoretically principled account of the phenomena available. Conceptions are idealized, quantitative relations are approximated or even transformed into qualitative relations, and information is represented in ways that exploit our visual and spatial capacities […] 'Explanatory power' appears to be shorthand terminology that may indicate any among a vast array of distinct properties that facilitate our ability to reason with and by our theories. (Woody 2004: 34-35)

Finally, the functional perspective highlights the necessity for philosophers of science to take into greater consideration social epistemological questions by 'stressing the ways in which knowledge production in the modern sciences is deeply, and inherently, social' (ibid: 85). From this perspective, explanations and explanatory discourse behave quite differently. An individual explanation can be detached from context and is often presented as static and isolated. But explanatory discourse is dynamic and its content and significance shifts as a scientific community evolves and new explanatory aims are developed. One implication of this is that if we are to make sense of explanatory scientific understanding as knowledge of explanations, the issue becomes less about treating the subject of understanding as an ideal epistemic agent directed towards isolated propositional and explanatory content, and more about paying attention to how that understanding is mediated by social mechanisms and institutional structures, and conditioned in the light of encultured, community standards of intelligibility

Therefore, the functional perspective offers one way of illustrating what is missing from explanationist and manipulationist accounts of explanatory understanding. It suggests that the conceptual significance of explanations in scientific practice is not exhausted by the knowledge that they convey or the psychological abilities of individual researchers; instead, what makes something explanatory is its capacity to coordinate ongoing inquiry and to sculpt and codify norms of intelligibility by which practitioners within a given research community make sense of and understand scientific phenomena. Explanatory understanding within the sciences is not just a matter of possessing knowledge of particular explanations of phenomena,

but involves being responsive to the ways in which scientific explanations communicate and constitute such norms of intelligibility. Explanations and explanatory discourse more broadly are thus integral to constituting the standards by which we understand domains of entities. The Ideal Gas Law, as we have seen, is explanatory not simply because it describes relations between gas phenomena, but because it 'instructs scientists how to think about gases *as gases*' (Woody 2004: 21). Moreover, this conception of gases orients scientists within the field of chemistry: it provides and determines the grounds upon which other chemical phenomena are intelligible and understood.

As such, from the perspective of the functional account, scientific understanding should not be thought of as a merely psychological or subjective phenomenon, idiosyncratic to individuals. Rather, it is importantly responsive to the norms embedded within social, scientific practices and the mutual interaction and collaboration of scientists within disciplinary communities. As Sabina Leonelli writes,

> It is […] important to keep in mind that scientific understanding is not the result of mere individual introspection. Its features are shaped by the necessity of intersubjective communication. The understanding acquired by one researcher will not be accepted as a contribution to science unless he or she is able to communicate insight to his or her peers, so as to make it vulnerable to public scrutiny and evaluation. Individual understanding becomes scientific only when it is shared with others, thus contributing to the growth of scientific knowledge and partaking in the rules, values, and goals characterizing scientific research. (Leonelli 2009a: 199)

If we consider scientific explanation in terms of functional role, then the construction and distribution of explanations within a scientific community is a means for coordinating scientific work in relation to such norms, standards, and shared values.

In summary, the functional perspective asks us to approach explanations in terms of how certain pieces of information and patterns of making sense of phenomena are treated as 'explanatory' within scientific communities and practices. It highlights the ways such explanatory patterns function to coordinate and guide coherent scientific practice. By stressing the functionality of explanatory discourse, rather than focusing solely on individual explanations and an analysis of the *explanans-explanandum* relation, the functional perspective takes a deflationary approach towards the search for adequacy conditions, but, in turn, nevertheless seeks to accommodate and explain why the standard models of explanation can be seen to be picking up on important features of why certain patterns are treated as explanatory. Furthermore, and most significantly, those explanations that are endorsed by a community no longer function just as repositories of accepted propositional knowledge about the world, but instead function to actively sculpt and perpetuate norms of intelligibility. Explanatory discourse has an essential social role in science, by enforcing communal norms

regarding 'what sorts of information are to be considered intelligible…and the types of reasoning that are legitimate within the community' (ibid: 86). As a result, the practice and scope of explanation is a thoroughly social affair, maintaining coherence and cohesion within and across scientific communities.

4 Explanatory Understanding, Skills and Normativity

Current accounts of explanatory understanding have tended to focus on the conditions that individuals must meet or abilities they must possess if they are to qualify as understanding why something is the case. I have so far challenged the focus on the individual subject by arguing that both the explanationists and manipulationists fail to pay attention to the irreducibly social nature of explanatory discourse in scientific practice, and by extension, the social nature of understanding.

The second point of contention for the current debate between explanationist and manipulationists is whether certain skills and abilities are necessary constituents of explanatory understanding, and if so, what form they take. Explanationists hold that the only thing relevant here is our knowledge of explanations. Manipulationists claim, by contrast, that we need to appeal to other factors to account for explanatory understanding. They identify certain cognitive abilities as the requisite feature for distinguishing between knowing an explanation and understanding something. So far, the main contenders are deductive reasoning skills, 'the ability to construct deductive arguments from the available knowledge' (de Regt 2009a: 26), and counterfactual reasoning; the ability of individuals to see how changes to one variable in an explanation will affect the other variables (Grimm 2010). The challenge to accounts of understanding which emphasise the role of skills and abilities has been forcefully put by Khalifa (2012), who argues that (i) with respect to de Regt, the appeal to deductive reasoning abilities is trivial, since no explanationist account would deny that individuals must be able to deduce the *explanandum* from the *explanans*, and (ii) with respect to Grimm, that the focus on counterfactual reasoning is already a central component of the explanation literature, as illustrated in Woodward's (2003) manipulationist account. The basic thrust of the challenge is that current accounts of skills and abilities and their role in constituting explanatory understanding are either trivial or unnecessary.

In this section, I challenge this claim by appealing to John Haugeland's (1998, 2007, 2013) account of the normativity of scientific understanding. One of Haugeland's central claims is that scientific understanding requires a particular kind of skill or know-how, which is the ability to tell when phenomena are not in accord with or violate the constitutive standards of a scientific domain of entities. Like Woody, Haugeland holds that scientific entities only

become intelligible in the light of these constitutive standards, paradigmatically in the form of scientific laws. I thus claim that Woody's and Haugeland's accounts complement one another: they both show that the connection between explanation and understanding operates at the level of the norms of intelligibility governing scientific practice.

Haugeland argues that there are four normative components to understanding entities within an intelligible domain. To illustrate, consider his primary example: a game of chess. Whether something is a rook or not does not depend on its physical form, but on its function assigned by the rules of chess. Haugeland argues that there are no rooks apart from the rules of chess. To be a rook means to be in accord with the constitutive standards of the game of chess. Similarly, these rules are not optional: not to move in the appropriate ways within a game of chess is to not to be a rook. Furthermore, the rules of chess are not just a shorthand for prescribing how players should behave or what they should do with pieces. Rather, the *constitutive standards* apply both to the players and to the pieces. They are all, in that respect, chess phenomena (Haugeland 1998: 320).

Yet, as Haugeland recognises, the fact that these constitutive standards govern what it means for something to be a chess phenomenon suggests the notion of an illegal move is problematic:

> The very idea of an illegal move flirts with paradox. That there should be chess phenomena at all *presupposes* that they accord with the rules (standards) that constitute chess as such. Nothing is so much as intelligible, let alone recognizable, as a chess phenomenon outside of the domain constituted according to those standards. […] Hence, strictly speaking, illegal moves are impossible. Yet they must be recognizable – and hence (in at least that sense) something that conceivably '*could*' occur – if the ruling-out is to be nonvacuous. That's the paradox. (ibid: 332)

The issue then is how would it be possible for something to even be intelligible if it is illegal; that is, if the constitutive standards of the domain mean that anything not conforming to those standards would be unintelligible? Haugeland's solution to this apparent paradox is to introduce a distinction between two types of skills: *mundane* and *constitutive* skills. Whereas '*mundane* skills…are the resilient abilities to recognize, manipulate, and otherwise cope with phenomena within the game… in effect, the ability to engage in play', 'a *constitutive* skill is a resilient ability to tell whether the phenomena governed by some constitutive standard are, in fact, in accord with that standard' – if they are, that is, behaving as they should (ibid: 323). The distinction between these two skills makes the recognition of illegal moves possible, Haugeland thinks, because it accounts for how the ability to recognise pieces and moves extends beyond the space of legal moves. A player must be able to recognise illegal moves, correct them, and insist on their illegality if performed by the other player. In other words, if mundane skills are the ability to play the game, to recognise pieces and moves insofar as they

constituted by the standards, constitutive skills are required so that players can tell when something has gone wrong (ibid: 334). Furthermore, although the two types of skills may, in practice, be hard to tell apart, there is a difference between the two is important. Haugeland points out that the constitutive skills 'cannot be just further mundane skills. They must be 'meta' or 'monitoring' skills vis-à-vis the results of mundane performances; for their essential exercise is to watch out for incompatibilities among those results' (ibid: 335). Finally, Haugeland argues that the authority and force of the rules of chess ultimately depends on a *constitutive commitment* of players to uphold and enforce them, in the sense of being intolerant of illegal moves. More substantially, however, this commitment is not just any other socially obligated commitment; rather, for Haugeland, it has a distinctively *existential* character:

> […] commitment, by contrast, is no sort of obligation but something more like a dedicated or even a devoted way of living: a determination to maintain and carry on […] Such commitment is not 'to' other players or people, or even to oneself, but rather to an ongoing, concrete game, project, or life. Thus, it is no more a psychological or an intentional state than it is a communal status; rather it is […] a way that relies and is prepared to insist on that which is constitutive of its own possibility, the conditions of its intelligibility. (ibid: 341)

This constitutive commitment thus underpins the responsibility on the part of at least some participants within a domain to holds its entire practice accountable to the constitutive standards by not tolerating violations of them.

There are therefore four components to Haugeland's account of what it takes to understand entities or phenomena within a domain: *constitutive standards*, *mundane* and *constitutive skills*, and *constitutive commitment* to uphold the standards throughout the domain. Each of these components are normative. Standards prescribe what is possible or not possible, and skills can be performed correctly or incorrectly. With respect to the sciences, Haugeland primarily identifies laws as examples of domain-constitutive standards, for 'laws are nothing other than true empirical universal claims with *modal* force: statements of what is and isn't possible among the phenomena' (ibid: 353). Like Woody's account of the Ideal Gas Law and its explanatory function, laws do not just describe relations between phenomena. They prescribe what is possible or not possible for the phenomena within a domain and as such govern scientific inquiry. Similarly, Haugeland draws on Kuhn's discussion of John Dalton in *Structure of Scientific Revolutions*:

> For Dalton, any reaction in which the ingredients did not enter in fixed proportion was *ipso facto* not a purely chemical process. A law that experiment could not have established before Dalton's work, became, once that work was accepted, a constitutive principle that no single set of chemical measurements could have upset. (Kuhn 1996: 133)

In other words, 'the principle of fixed proportions became a constitutive standard for what it is to *be* a chemical reaction – and thereby also for being an element or a compound' (Haugeland 1998: 258).

Constitutive skills presuppose the mundane skills because without them scientific research would not proceed effectively. In this respect, mundane skills pertain to the familiar scientific know-how and abilities that scientists acquire through training and experience. For example, Sabina Leonelli (2009a) offers a useful classification of the types of 'mundane' skills that are required in the biological sciences.[12] First, there are *theoretical skills*, which 'involve the mastering the use of concepts, theories, and abstract models' (Leonelli 2009a: 201). From this perspective, the kinds of skills that de Regt and Grimm identify would be specific types of theoretical skills (deductive and counterfactual reasoning abilities of theoretical knowledge). Second, Leonelli argues that there are *performative skills*, which involve the familiar, practical know-how that underpins scientific practices. Performative skills are acquired through training, education, and through direct intervention with the material resources of a particular discipline. Such skills are developed through repeated practice and echo Dreyfus' notion of 'skilful coping' as the cultivation of and adaption of embodied abilities to the demands of the environmental circumstances in which the scientist is placed. Third, there are *social skills*, which refer to the 'ability to conform to existing standards…[and] denoting the ability of researchers to behave and express their insights in ways that are recognized by their peers and/or other participants in their social context' (ibid: 201). Constitutive skills are thus further skills that monitor the successful execution of these mundane skills.

Importantly, Haugeland argues that skills must be *reliable* and *resilient*. Reliability concerns the familiar requirement of scientists to carry out procedures in a consistent fashion in ways that are in accord with generally accepted epistemic norms and standards. Resilience, on the other hand, designates the ability to respond in the appropriate manner to apparent breakdowns, failures of reliability or incompatibilities between the results of mundane skills.

For Haugeland, laws *qua* constitutive standards, along with constitutive and mundane skills, ultimately presuppose commitment. In scientific practice, commitment refers to scientists' capacity to uphold the standards of specific domains of entities by not tolerating violations of them. There are three dimensions to scientific commitment. At a basic level, scientists must be committed to exercise the relevant mundane and constitutive skills in the right kind of way and to exercise them in accord with the norms of disciplines or communities. Equipment must be used correctly, theoretical calculation, modelling and analysis must be performed in the right ways, experiments must be designed and implemented properly, and so

[12] 'Mundane' is not meant in a pejorative sense. As Leonelli's examples show, scientific research presupposes a complex variety of technical and cognitive skills in order to function successfully.

forth. At this stage, the focus is on making sure that these skills are exercised in the right way, so that what appears through observation, detection, and analysis genuinely manifests the entities or phenomena in the domain, rather than being the product of human error.

However, a second level of commitment is important when all these skills and procedures have been carried out appropriately but nevertheless still yield results that are incompatible with the domain's constitutive standards. At this level, the commitment is to adjust and refine the skills and standards in order to rectify the problematic results. If, for example, a perfectly executed procedure yields a different outcome when carried out by another scientist, then that procedure may require repair or revision. Ultimately, what matters is the extent to which those procedures yield result that are compatible with the constitutive standards. The *resilient* nature of skills becomes important here. Haugeland argues that skills and standards cannot be revised on a whim; they must be able to withstand the possible recalcitrance of entities to a certain extent: 'If performances or skills are "revised" or "repaired" casually, at the first sign of trouble, then nothing is seriously excluded, and all "testing" is a farce' (Haugeland 1998: 334). The third level of commitment is required at the point when repeated and persistent errors or incompatibilities do not go away even after procedures have been checked, modified or substantially replaced. At this point, Haugeland argues that scientists must be willing to abandon an entire domain of research. The Kuhnian overtones are palpable here: ultimately, it is the commitment of scientists to abandon research programmes if they are no longer accountable to how the world is that sustains the normativity of genuine scientific understanding.

Each of these elements is essential for scientific understanding on Haugeland's analysis. In this respect, Haugeland's account of scientific understanding represents an important extension of Woody's functional perspective on explanation. Woody's analysis of the Ideal Gas Law emphasises its domain-constitutive normativity, but Haugeland's arguments show that its explanatory function can only facilitate scientific understanding in the light of more general skills and commitments. In this respect, explanatory understanding is not simply a matter of possessing certain knowledge. Rather, if understanding involves grasping entities within the light of the constitutive standards of a scientific domain, then to explain something is to 'show that and how something can be understood' (ibid: 260). Within scientific practice, the social function of explanatory discourse is to show how entities discovered or studied within that domain are in accord with its constitutive standards: 'to explain is to make intelligible…to show that, given the laws and the actual concrete circumstances, some actual phenomenon *had* to occur – or, in some cases, even just that it *could* occur is to explain it' (ibid: 353). In this respect, Haugeland and Woody reach a similar conclusion from opposite directions. Woody's analysis of the function of explanation in scientific practice highlights its central role in constituting and perpetuating the norms of intelligibility or constitutive

standards for domains of entities: 'explanations encode the aims and values of particular scientific communities, telling practitioners what they should want to know about the world and how they should reason to get there' (Woody 2015: 81). Similarly, Haugeland argues that the function of explanations is not just to impart explanatory content or information, but to show how the phenomena exhibited in an explanation are intelligible in the light of the constitutive standards of a scientific domain. From this perspective, explanatory understanding is fundamentally a matter of grasping entities and phenomena in the light of constitutive norms and standards of intelligibility. As Haugeland puts it:

> Phenomena […] are *understood* insofar as they are recognized as being in accord with the standards constitutive for their domain. If understanding is lord of the domain, explanation is concierge: to *explain* a phenomenon is to 'show it in', to exhibit its allowability according to the standards, to let it be. (Haugeland 1998: 286).

Furthermore, Woody argues that connecting explanation and understanding at the level of normative intelligibility and attending to the function of explanation in scientific communities allows us to recast the significance of the standard accounts of scientific explanation. In Haugeland's terms, the various models of explanation make explicit different possible ways in which it is possible to show how phenomena can be shown to be in accord with the norms of intelligibility of a scientific community. For example, the aspiration of unificationist accounts of scientific explanation (Friedman 1974; Kitcher 1989), which takes the explanatory status of theoretical frameworks to be a consequence of the unification they supply, has often struggled with (a) articulating reasons why such a unifying framework would be desirable in scientific contexts, and (b) justifying the presumption of metaphysical orderliness that would explain such unification. As Woody argues, the justification for the preference for unification is methodological rather than metaphysical: it is 'built …on the conditions for coherent cognitive labour' and 'reflects the nature of coordinated intellectual work rather than the nature of the phenomena of study' (Woody 2015: 85). Moreover, the fact that both the unificationist and the deductive-nomological account treat individual explanations as arguments (recall that both are *epistemic* approaches to explanation), is accommodated within the functional perspective because it treats argumentation as a central tool in the process of accepting certain explanations as intellectually binding for a community.[13] Similarly, Woodward's causal-interventionist account, which highlights the role of addressing 'what if things had been different questions', is captured by the normative function of scientific

[13] The argumentative structure of scientific explanation is also foregrounded in recent approaches that draw on argumentation theory and rhetorical theory for understanding scientific explanation. See Rehg (2009), Faye (2014), and de Donato-Rodriguez and Zamora-Bonilla (2009b).

explanation because such explanations contribute to the delimitation and articulation of what is possible or not possible within a domain.

5 Concluding Remarks

This chapter has analysed explanatory understanding at the level of norm-governed scientific practices. I began by discussing competing accounts of explanatory understanding in the current literature and identified two central issues. First, current accounts have presumed that the appropriate level of analysis of the connection between explanation and understanding is at the level of the individual subject. I argued that this individualistic focus obscures the social dimension of explanatory understanding in scientific practice. Second, it is unclear what skills or abilities are required for explanatory understanding, if any at all. De Regt's and Grimm's accounts of skills were problematic because, as Khalifa argued, they risked being either trivial or unnecessary.

This chapter has offered arguments that address both issues and show how their resolution lies at the level of norm-governed, social scientific practices. I first examined Woody's functional perspective on explanation, which emphasises the social nature of explanatory discourse in scientific practice. For Woody, explanations serve to constitute, sculpt and inculcate the norms of intelligibility of scientific communities and disciplines. As such, they play a significant conceptual role in guiding how to think about entities within their domain. Second, I drew upon Haugeland to show how scientific understanding is similarly responsive to scientific norms of intelligibility, or more broadly, domain-constitutive standards. Therefore, both accounts emphasise the social normativity of explanatory understanding. I have thus sought to show how explanatory understanding can be made sense of at the level of norm-governed social practices. From this perspective, understanding and explaining both operate in the light of the norms and standards of intelligibility that govern scientific domains. This supports the broader claims of this thesis: that understanding is not just a matter of possessing the right kind of explanatory knowledge, but fundamentally concerns the interconnections between understanding, normativity and scientific practice.

Conclusion

Contemporary philosophical work on the concept of understanding has provided rich and detailed analyses of what it means to understand. I have sought to contribute to these analyses by arguing that human understanding is fundamentally connected to normativity. Human understanding is neither just a contingent psychological state or ability, nor simply the possession of the right kind of knowledge about the world. Instead, I have argued that to understand something means to understand it in the light of the norms of intelligibility that govern our social practices. I have shown that this connection between understanding and normativity is equally constitutive of scientific understanding. By analysing the use of models and scientific explanations in scientific practice, I have argued that each facilitates scientific understanding through its crucial role in conceptually articulating a 'scientific space of reasons'. From this perspective, scientific understanding fundamentally concerns our ability to be responsive to the conceptual norms guiding our research practices.

In this conclusion, I will first outline the main claims of each chapter and show in what way they have contributed to and supported the main claim of this thesis. I will then consider the broader philosophical implications of the approach I have taken to the concept of understanding, before identifying three avenues for further research. These suggestions for further research are natural extensions of the central claims made in this thesis, and I will indicate in what way they may bear on the contemporary debates.

1 Summary of the Thesis

Any systematic study, in philosophy or any other form of inquiry, begins from certain assumptions about its subject matter. These assumptions will guide how we make sense of the relevant questions, frame which problems are relevant or worth investigating, and suggest which solutions will be treated as appropriate or acceptable given the research problem. We will often try to make explicit the assumptions that inform our approach to a subject matter, but setting out what we take for granted can be challenging. Philosophical accounts of the concept of understanding are no different. In Chapter One, I argued that how we think about what it means to be a subject that is capable of understanding will have important ramifications for how we then approach the concept of understanding itself. Within the contemporary

debates, most accounts have proceeded by assuming the subject of understanding to be the standard subject of traditional epistemology: the ideal epistemic agent, possessor of coherent, justified true beliefs, approximately true or correct explanations, and the right kinds of mental machinery and processes. I do not want to suggest that this is fundamentally inadequate as an approach; it has undoubtedly shone some light on how as subjects we understand things in the world. However, I do contend that exclusively taking this approach risks overlooking important features about what it means to be an epistemic agent and what it means for such an agent to understand. As such, Part I of this thesis addressed the nature of the subject of understanding and in Part II, I focused on the vehicles of scientific understanding in the light of this conception.

In Chapter One, I proposed an alternative starting point, one that began with Heidegger's phenomenological analysis of the subject of understanding in *Being and Time*. Phenomenological conceptions of subjectivity typically treat the subject not as an ideal, rational knower, but as an embodied agent, who is embedded within a meaningful and intelligible world. From a Heideggerian perspective, I argued that understanding is first and foremost a practical, normative capacity to deal with one's surroundings and to engage with entities in terms of their intelligible possibilities. As a subject of understanding - as Dasein - I encounter and make sense of things in the world by holding them to (often not fully explicit) norms or standards that are embedded within our social practices and contexts of activity. These norms or standards are not primarily epistemic, moral or political norms, but rather more basic norms of intelligibility. Without these norms, Heidegger argues, things in the world would lack intelligibility or meaning; they would, ultimately, lack 'being'.

Heidegger argues that the norms and standards by which we make sense of entities and in terms of which we understand them, is only possible if we understand ourselves as being something capable of success or failure. I argued that, for Heidegger, I must not only be capable of responding to the norms embedded within our social practices, I must ultimately be able to make sense of the norms as norms, as having an authority and force over what I myself say and do. In Kantian jargon, understanding things in the world requires more than acting in accord with a norm: it requires acting for the sake of a norm. The possibility of acting for the sake of a norm is an essential constituent of Dasein's self-understanding: the norms within a social practice only take on normative force for me if I can understand myself as someone trying to be what those norms require. The central claim of Chapter One, therefore, is that as a subject of understanding, I not only understand things in the light of the norms governing our social practices, I am also capable of responding to those norms as norms; to be able to question, that is, whether and how those norms have authority for me with respect to the kind of person I am trying to be.

In Chapter Two, this claim was developed and expanded upon by drawing upon Brandom's 'normative inferentialism'. It is often argued within contemporary accounts of understanding that understanding requires more than holding particular beliefs about the world. A key claim here is that understanding involves inferential abilities: our ability to grasp or see the connections between our beliefs within a body of knowledge, to be able to reason counterfactually, and to have a sense of what follows from what. I argued that Brandom's normative inferentialism represents an important qualification of this idea by holding that the inferential connections between beliefs only acquire content because they take on normative significance within what Brandom calls the 'game of giving and asking for reasons'. The inferential connections between our beliefs have normative import insofar as they are constituted by relations between the different kinds of normative statuses we hold within our social practices and community. For Brandom, conceptual understanding is not merely a matter of knowing those inferential connections, but of grasping which connections are appropriate with respect to the rules and norms of material inference. The conceptual normativity governing our inferential abilities is two-dimensional: it pertains to understanding not only the appropriate circumstances in which concepts can be applied, but also the appropriate consequences of their application. However, crucially, for Brandom, the norms governing the inferences from certain circumstances to particular consequences are not fixed independently of our practices. How and where a concept applies, and what follows from it, is not simply a matter of our psychological abilities, but is rather worked out in a continual process of determining which inferences as a community we should endorse and which we should revise or abandon. The nature of conceptual normativity can thus be seen as a natural extension of the central claims I took from Heidegger: not only do we understand things in the light of particular (inferential) norms that govern our practices, we are continually responsive to those norms as norms, of working out which ones we ought to endorse and take as binding for a practice or community.

The focus of Chapter Three was on the notion of normativity itself and how it informs our conception of human, social practices. I presented arguments against two conceptions of norms: regulism, which treats norms as explicit rules, and regularism, which treats norms as regularities in our behaviour. I argued that both conceptions only provided the semblance of normativity. The alternative suggestion was made that norms should be thought of as clusters of normative attitudes. On this conception, normative attitudes are sustained in the mutual interaction between participants within a practice by holding each other to account. In that respect, norms do not exist independently of our practices and only have authority within our lives insofar as we continue to take them to be binding. Drawing on Rouse's normative conception of practices, I argued that the normativity of human practices is sustained by their temporally extended nature and in the fact that there is something at issue and at stake in

continuing that practice. What is at issue is how and whether the practice should continue and what is at stake is how those issues are resolved in the continuation of the practice. On this conception, normativity is an irreducible feature of human practices, which is located first and foremost in the attitudes of mutual accountability that we take to one another, but more generally in what is at issue and at stake in a practice continuing in one way rather than another. To reiterate the familiar theme, what matters for the normative structure of a practice is that as agents embedded in such practices, we are the kinds of agents who can respond to norms as norms.

The first three chapters therefore focused on how and in what sense understanding is connected to normativity from the perspective of the epistemic agent. First and foremost, understanding refers to our ability to understand things in the light of norms and standards. I understand something not just because I have certain knowledge about it or because of certain psychological abilities, but because I already have a grasp of what is possible or not possible for the entities with which I am dealing; that is, in what sense they are intelligible at all. However, I have also argued that to be a subject of understanding means to be capable of responding to norms *as* norms. For Heidegger, this is primarily evident in individual cases: which norms I take as binding upon me in terms of the person I want to be. For Brandom, the point applies more generally to the kinds of inferential norms of material inference that we as a social community endorse. Finally, this issue of normative reflection and endorsement lies at the root of a normative conception of human practices, which are sustained by a continual issue of what it means for that practice to be at all, in the sense of how and in what way it provides an intelligible articulation of the world.

Chapters Four and Five develop these claims with respect to scientific understanding. Philosophers of science have been concerned less with the nature of the epistemic agent who understands, and more with how the various elements of scientific inquiry facilitate and provide scientific understanding. These chapters showed how some vehicles of scientific understanding, specifically models and explanations, facilitate that understanding through their role in articulating how scientific phenomena and objects of research become intelligible in the first place.

Chapter Four focused on how scientific understanding is facilitated using models. One of the key epistemological issues for current accounts of model-based understanding concerns the truth-status, or factivity, of the understanding provided by models. The use of idealized assumptions, fictions, and deliberate falsehoods in the construction of models would seem to challenge the idea that scientific understanding should be factive. I argued, following Henk de Regt and Catherine Elgin, that truth and factivity is not of primary importance for making sense of how models provide scientific understanding. While not denying that our understanding should be accountable to the world, I nevertheless argued that the ability to

assess the truth or representational accuracy of the claims derived from models presupposes that what a model says about the world is already intelligible. The issue for model-based understanding is working out how and in what sense models present their targets intelligibly in the first place. To do this, I argued that models should be seen primarily as tools for conceptual articulation, which I expanded upon by drawing on Rouse and Kuhn. By combining a view of models as epistemic tools with a focus on their inferential properties, I argued that models conceptually articulate the world in (at least) three ways. First, models place phenomena into the space of reasons by presenting them as things that can be assessed, reasoned about, and investigated further. Second, conceptual articulation affords modal understanding, since models are not only used for determining the way things actually are, but also for articulating the kind of ways in which phenomena could be. I argued that this should be seen to be of equal importance in scientific research, insofar as it discloses a space of intelligible possibilities in terms of which phenomena can make sense.

Third, most crucially, models play a key role in establishing the conceptual normativity of a domain of entities. The normative dimension of modelling pertains to whether they say anything intelligible about the world at all. I argued that the possibility of this is at stake in the way in which models can articulate projectible patterns concerning their target phenomena that are then treated as candidates for endorsement within a scientific community, field or discipline. The point, here, is not that models automatically provide us with claims about phenomena that we ought to accept. Rather, the issue is whether scientists take the claims derived from model-based analysis and reasoning to extend beyond the domain of their own immediate application and thus be worthwhile exploring further. This normative dimension to models also pointed to the connection to understanding, which, I argued, drawing on Haugeland's analysis of pattern recognition. Pattern recognition constitutes the ability to tell, on the one hand, whether a pattern is there at all, and on the other hand, to be able to tell if a particular case is an element in that larger pattern. It is only at this level, I suggested, that questions concerning the truth or representational accuracy of model-based understanding come into force, because then the existence of a pattern in the phenomena, or whether a particular case counts as an element of that pattern, can become subject to procedures of testing and empirical assessment.

In Chapter Five, I argued that explanatory understanding cannot be divorced from the constitutive standards that govern domains of entities and the communities and disciplines that study them. I argued that current accounts of explanatory understanding have failed to take seriously the way scientific explanation functions in scientific practice. Andrea Woody's recent functional perspective on scientific explanation, which emphasises the essentially *social* nature of explanatory discourse, argues that explanatory discourse serves to sculpt and inculcate the norms of intelligibility for scientific communities and disciplines. Established

explanations constitute endorsed patterns of reasoning for scientists in their fields. Exemplary *explananda*, such as the Ideal Gas Law, perform crucial conceptual work in the field of chemistry in not only describing how particular gas phenomena are related, but more fundamentally guiding scientists how to think about gases *as gases*.

In this respect, I argued that the Ideal Gas Law can be thought of as a special case of Haugeland's more general analysis of domain-constitutive standards. To be a phenomenon or entity within a particular domain is to conform to these standards, or within levels of tolerable divergence. Haugeland's arguments show how scientific understanding operates in relation to these constitutive standards through the application of scientific skills, which constitute both our capacity for skilful performance within scientific practice but also for recognising possible violations of the constitutive standards. At the root of this interconnected normative structure of scientific understanding is the notion of constitutive commitment. I argued that commitment is essential for upholding constitutive standards by not tolerating violations of them, repairing or revising our skills and standards if they yield repeated violations. Ultimately, however, it is a commitment to abandon the field of research if its norms and standards fail to tell us anything about the world. The commitment to repair, revise or abandon the complex system of norms that scientific communities, disciplines and fields construct is at the root of how scientific understanding can be accountable to the world. From this perspective, understanding and explanation were shown to be two kinds of the same coin: if understanding is grasping entities or phenomena in the light of a domain's constitutive norms and standards, explanation is showing how those entities are intelligible and in accord with those standards. In that respect, explanation and understanding mutually reinforce each other, but they can come apart. Thus, pace explanationist accounts, understanding is not reducible to knowledge of explanations insofar as explanations only work if we already have some grasp of what is possible or not possible for entities in an explanatory domain; that is, if understanding is already in place.

In sum, this thesis has defended the claim that understanding is fundamentally a matter of being responsive to and responsible for the norms of intelligibility governing our social practices.

2 Further Research

The primary focus of this thesis has been to analyse the connection between understanding, normativity, and meaning. To do this, I have drawn on different schools of thought; more broadly, however, I have argued that these approaches share common concerns, despite significant terminological differences. In this respect, this thesis will have achieved its broader

aim if it has demonstrated that not only can phenomenology, Pittsburgh neo-pragmatism, and philosophy of science be fruitfully brought to bear upon one another, but also that they can be productively applied to general issues in contemporary epistemology and philosophy of science.

Characterising understanding in general, and scientific understanding in particular, as being responsive to the normative landscape of human practices offers a different set of conceptual resources for making sense of understanding. In what follows, I identify three areas of research that would develop the claim defended in this thesis: (i) understanding and objectivity; (ii) the social nature of understanding; and (iii) understanding, commitment and responsibility.

2.1 Objective Understanding

I have argued that understanding is first and foremost directed towards the intelligibility of things. One of the main challenges faced by any account of understanding that emphasises intelligibility is that it makes understanding an entirely 'internal affair' (Grimm 2010: 90) that sacrifices its fidelity to the facts and the truth of the matter. As such, the account I have offered may seem orthogonal to the objectivity required of scientific understanding. The trouble is that whether something is intelligible or not seems to be merely a subjective affair; it is a value which we project onto the world, but it offers no guarantee that that is how the world really is. Furthermore, this worry is often exacerbated by the possibility that false theories can often be more intelligible than correct ones, particularly to scientists within different historical periods. For example, Wilkenfeld (2017) suggests that in the sixteenth century, Ptolemaic astronomers had good reasons not to immediately accept the Copernican heliocentric model because its predictions were less accurate than those of the Ptolemaic alternative. As such, the Ptolemaic model was taken to be more intelligible. The problem with this, Wilkenfeld contends, is that it seems to imply that Ptolemaic astronomers not only understood why the planets exhibit retrograde motion, but moreover, 'that they actually understood better than the Copernicans who correctly put the sun at the center of the solar system' (Wilkenfeld 2017: 1278). So, intelligibility is not sufficient for understanding. In response, various commentators have argued that genuine understanding requires representational accuracy, that it should be factive, and that understanding should be objective. As Grimm argues, genuine understanding should move beyond simply what is intelligible to the subject, or what they are capable of grasping. Rather, objective understanding is

> […] the kind of understanding that comes not just from grasping a representation of the world that fits with one's world picture, but also from grasping a (more or less) correct representation

> of the world. Objective understanding therefore entails subjective understanding but goes beyond it, requiring that the grasped representation in fact obtains. (Grimm 2010: 91)

Although my own approach to understanding has generally prioritised conceptual meaning and intelligibility, I have sought to show that just because understanding takes place in the light of norm-governed, social practices, this does not mean that it cannot be accountable to how the world is. My own view, as discussed in Chapters Four and Five, is that the possibility of determining whether our claims about the world are representationally accurate, true, or factive, depends on their prior intelligibility because it is only when such claims are construed as meaningful that they can become candidates for assessment and testing. I do not want to deny, in that sense, that objective understanding is at the very least a regulative ideal for scientific understanding. Nevertheless, I think that simply stating that truth and representational accuracy is a condition for understanding only names the problem; it does not solve it. Furthermore, clarifying the objective nature of understanding is complicated by several further issues, which I think only become visible once we recognise the connection between understanding and conceptual normativity.

First, the claim that understanding should be accountable to the world becomes significantly more complex in the light of Sellars' critique of the so-called myth of the given. Sellars provided a powerful argument against the idea that the world could have any rational bearing on our understanding unless it was already conceptually articulated, and thus already intelligible. As McDowell (1994) argued, is that this makes it impossible to take a 'sideways on' view of how our concepts and understanding fit with how the world objectively is, since any assessment already presupposes our rational second-nature. If this is the case, then securing objective understanding is not something that can be decided from the 'outside', as it were, but must be something that can only be done from 'within'. This point extends to scientific inquiry too. Scientific phenomena, for example, are never just 'given' to us; rather, they require complex arrangements of instruments, equipment and technical know-how in order to show themselves in an intelligible way.

In a similar vein, Catherine Elgin (2009) has argued that representational accuracy is itself not enough for scientific understanding. Instead, Elgin argues that we need to distinguish between accuracy and objectivity: 'A hunch may be accurate. My completely uninformed guess as to who will win the football game may turn out to be correct. But there is no reason to believe it, for it is entirely subjective' (Elgin 2009: 88). Her point is that the objectivity of our representations of the world turn not on their representational accuracy, but on their relation to reasons:

> If a representation is objective, it is assessable by reference to intersubjectively available and evaluable reasons, where a reason is a consideration favouring a contention that other members

> of the community cannot responsibly reject […] scientific objectivity involves answerability to the standards of a community. According to these standards, among the factors that make a scientific result objective are being grounded in evidence, verifiable by further testing, corroborated by other scientists, consistent with other findings, and delivered by methods that have been validated. (ibid: 88-89)

Elgin's point can thus be read as complementing the position I set out in Chapter Four. What matters for scientific understanding is that models help to articulate claims about the world that can be candidates for endorsement within a scientific community, which means that it can be subject to the kinds of scientific testing and empirical assessment that Elgin describes. Furthermore, I argued in Chapter Five that explanations provide understanding by showing how the phenomena they are purportedly about accord with the constitutive standards of the community. Another way of reading this would be to say that explanations provide us with deeper reasons for why we want might want to endorse the picture of the phenomenon that the explanation provides. From this perspective, the objectivity of understanding needs to be addressed in relation to the intersubjective standards of scientific communities. Furthermore, Haugeland's own account of how objective understanding is possible does not turn on the possibility of correct representations of the world but on the possibility of an 'excluded zone' for any domain of entities, which is made up of the set of entities that are recognisable but illegal given the constitutive standards of a domain. Objective scientific understanding becomes possible only if we can recognise entities that do not conform to the constitutive standards – in Haugeland's case, scientific laws – of a domain. Only then does the possibility of our skills, procedures, claims and standards being accountable to how the world is, arise.

Second, Lorraine Daston and Peter Galison (2010) have shown how our notion of objectivity actually incorporates several competing strands that do not neatly fit together.[1] Their social and conceptual history of objectivity shows different conceptions of objectivity having greater grip at different historical moments; indeed, in doing so, they argue that objectivity is itself a historically situated notion and ideal. According to their account, the eighteenth century was dominated by a conception of objective representation governed by the norm of 'truth to nature', the nineteenth century by 'mechanical objectivity', and the twentieth by 'trained judgement'. Importantly, these conceptions do not discretely supplant one another, but versions of each are entangled with one another through the course of the development of the sciences. Daston and Galison take a 'bottom up' approach to objectivity: rather than offering a philosophically articulated ideal, they purport to begin in historically evolving scientific practices. This invites consideration of what it means to have objective understanding, which presumably has itself varied across different periods within the history

[1] See Daston and Galison (1992, 2006), and Daston (1992).

of science. Although it has already been recognised that the standards for what counts as understanding have varied across the history of science (e.g., de Regt and Dieks 2005), Daston and Galison's approach invite closer scrutiny of how different conceptions of objectivity would inform our conception of scientific understanding.

The final consideration that I think should be brought to bear on our conception of objective understanding is that the sciences do not just seek any truths. Most truths about the world are irrelevant to scientific inquiry. In turn, the sciences do not seek representational accuracy for its own sake. Rather, the sciences seek significant truths, insightful truths, truths and objectively assessable claims that are relevant to the local and contextually determined epistemic aims and goals of inquiry. Furthermore, those claims should be conducive to further research: as we saw in Chapter Four, the normativity of modelling turns on whether the models articulate projectible patterns in the world that can be further explored and investigated. Recognising this claim would also highlight the need for further investigation into how understanding is connected to such normative assessments as significance, relevant, and salience.

2.2 The Social Nature of Understanding

In this thesis, I have stressed the social dimension of understanding. Our understanding is normatively accountable to other participants within a practice, as we navigate the game of giving and asking for reasons. I have thus argued that our understanding cannot be divorced from the public, intersubjectively available norms and standards that shape and structure our understanding of the world. Furthermore, I argued in Chapter Five that current accounts of explanatory understanding are inadequate because they have either excluded, or at least not adequately accounted for, the social dimension of explanation. Explanations do not exist in a vacuum; they are constructed and offered in specific circumstances, and they are shared between agents with social groups, communities, and disciplines.

I have characterised the social dimension of understanding in terms of our responsiveness to the norms governing our practices. However, there is much more that could be said. First, as I noted in Chapter Five, although we might recognise that social practices, contexts and norms inform an individual's understanding, what has yet to be investigated in the current debates on understanding is under what conditions we attribute understanding to groups, rather than individuals. This is particularly pertinent with respect to scientific understanding. Modern scientific inquiry is often distributed across several different communities, which would suggest that scientific understanding is embodied in communities of practitioners in highly complex and interconnected ways (Traweek 1988; Galison 1997). The conditions of

contemporary science, including the advent of 'big science' in which large groups of scientists coordinate their labour to conduct research, is carried out through multidisciplinary projects occurring within international networks (e.g., Shrum et al. 2007). Within social epistemology, the traditional conditions of knowledge have been supplemented with additional factors when we consider knowledge in a social setting. For example, questions of authority, testimony, shared values, and power relations take on a greater salience from this perspective, and one promising direction for further research would be to consider how these factors shape understanding.

Similarly, Adam Toon (2015) has drawn on current work in cognitive science, in particular research on the embodied and extended nature of cognition, to show that understanding itself can be treated as extended in several ways. One way to develop this line of inquiry would be to consider how ideas and concepts from work on distributed cognition (e.g., Hutchins 1995) could be brought to bear upon the social nature of understanding. Distributed cognition emphasises the necessity of coordination and collaboration within groups to achieve aims and goals which would not have been possible for individuals within groups working by themselves. From this perspective, there may be important differences in the conditions under which scientific understanding is obtained and attributed when we focus on the group level, rather than the individual.

Furthermore, I have argued that normative accountability is essential to our understanding. Only because we are held to account within our social practices does our conceptual understanding take on meaning and significance. Recently, Eric Winsberg, Bryce Huebner, and Rebecca Kukla have examined the way in which accountability is spread across massively distributed collaborative work. Questions about who exactly is accountable for scientific claims about the world, when there are significantly disparate goals of inquiry, methodological norms, or epistemic values across a wide range of different research groups, suggests a possible 'crisis of accountability' (Winsberg et al. 2014). Further research could clarify how accountability is mediated across different scientific groups and communities by focusing on who is responsible for providing reasons for claims in a massively distributed game of giving and asking for reasons.

2.3 Responsibility and Commitment

A final way to develop the claims defended in this thesis would be to focus in greater detail on the connections between understanding and responsibility. There are two kinds of responsibility that I discuss here: epistemic responsibility and existential responsibility.

With regards to epistemic responsibility, we have seen that, for Brandom, in asserting something, I not only signal my commitment to the content of that assertion, I also undertake a justificatory responsibility to provide reasons for my commitment if I am challenged. As a condition for understanding, I must be epistemically responsible; that is to say, I must make sure that my beliefs, the connections between them, and the justification I offer are managed in a proper and epistemically responsible way. The connection between responsibility and understanding in this sense thus points towards a form of internalism regarding understanding. Internalism about understanding would hold that an agent only has understanding if she is able to access and offer reasons in favour of her understanding. A strong form of internalism about understanding has been defended by Linda Zagzebski. According to Zagzebski, the epistemic value of understanding lies not with its factivity, but in what she calls its 'transparency'. She claims that, 'it may be possible to know without knowing one knows, but it is impossible to understand without understanding one understands' (Zagzebski 2001: 246). The idea here is that, to count as understanding the agent must have access to the fact that she understands.

Some authors have argued that internalism places unnecessarily demanding constraints on our understanding. In that respect, their arguments track familiar criticisms of internalism about knowing, which are often challenged for being overly intellectualist. Grimm, for example, attacks internalism about understanding. Grimm holds that understanding is knowledge of dependency relations, such as causal relations, and as such must meet the justification condition required for knowledge. However, he takes this to be primarily a matter of reliable belief formation, which need not necessarily be accessible to the subject. Thus, externalism about understanding, in a form of reliabilism, would be less intellectualist and demanding, without sacrificing some form of justification requirement.

From the perspective developed in this thesis, there is good reason to think that we cannot completely dispense of some form of internalism about understanding. Yet this internalism need not necessarily exclude external, reliable belief formation either. Both Sellars and Brandom, for example, argue that our ability to respond reliably to things in our environment is a necessary condition for understanding. Both hold, for instance, that to count as understanding, I must be able to reliably respond to things in my environment. Moreover, Brandom argues that the entitlement to the commitments I inherit from others within a social community need not be transparent to me: I may not be able to offer precise reasons for my commitments, but if I can defer to an appropriate source, I retain my claim to understanding. However, both Sellars and Brandom contend that reliability by itself is not sufficient for understanding. To count as understanding, the agent must have some minimal grasp of what follows from reliably formed beliefs, possible evidence for those beliefs, how they might be challenged, and what would be appropriate ways to meet those challenges.

This might seem to lead us straight back into a Zagzebskian strong form of internalism. But I think where Sellars and Brandom depart from strong forms of internalism is that they hold a less demanding conception of epistemic responsibility. To understand the basic difference in their position, we can follow Michael Williams (2008) in distinguishing between two senses of responsibility. On the one hand, responsibility captures the requirement that I manage my beliefs in an epistemically responsible way; that is, responsibility involves adherence to properly standards of epistemic conduct. On the other hand, I am responsible in the sense that I am accountable or answerable for what I say and do, and this accountability is meted out in terms of my capacity to offer reasons for my claims and actions if challenged (Williams 2008: 2).

The difference between the two can be illustrated by appealing to two different ways of thinking about epistemic justification. One way to think about the justification of our beliefs is that our personal justification is wholly dependent on prior, evidential grounds. On this approach, one is epistemically responsible in believing a given proposition only if one's belief is based on adequate, non-defeated evidence. Williams labels this the 'Prior Grounding Requirement' (Williams 2001: 24). On the basis of this requirement, I am not entitled to a belief unless there are good, evidential, prior grounds of which I am self-consciously aware. A rival model to this is Brandom's 'default and challenge' model of justification (MIE: 176-179). On this model, we can be justified in believing that-p without our having done anything specific to earn that entitlement. Instead, '[d]efault entitlements only accrue to accredited epistemic agents. The status of the epistemic subject, along with that of the moral agent, is earned through training and acculturation' (Williams 2008: 13). On the default and challenge model, my entitlement to a claim is assumed, and only comes into question if I make lapses in judgement or draw erroneous inferences. What matters is that, in being held accountable, I can then provide reasons for my claims; but, more substantially, the epistemic responsibility involved here is recognising that I have a responsibility to respond to demands for reasons or evidence for my commitments. In that sense, I am accountable in the sense I recognise that I should be prepared to give an account of my beliefs if called upon to do so.

The point therefore is that understanding epistemic responsibility qua accountability, which pays due acknowledgement to our normative status as agents, offers a possible route for reconciling internalism and externalism about understanding. For on this conception, we can grant that epistemic responsibility involves adherence to proper epistemic standards and concede that the processes for adhering to these need not be self-consciously transparent to the agent, while at the same time holding that I must recognise that my claim to being an epistemic agent who understands entails certain responsibilities in the sense that I am accountable for what I say or do. If this line of reasoning is plausible, it provides one way of

showing how the notion of responsibility is connected to understanding, while offering a possible way of reconciling certain positions within the current debates.

But there is a deeper way in responsibility is essential to understanding that would merit further research, which I shall call 'existential responsibility'. This existential form of responsibility can be found in Haugeland's work, but also, I would suggest, ultimately traces back to Heidegger. Haugeland, as we have seen, argues that the norms and standards inherent in our practices require an existential or constitutive commitment. For Haugeland, it is only because I take responsibility for the standards in the practices in which I am engaged, that those practices are possible at all. The fact that I take responsibility by upholding those norms or standards, making sense of things in the light of them, and not tolerating violations or divergences from those norms or standards, means that those practices have a normative grip on me and in that respect, differ from mere regularities of behaviour.

Haugeland thinks this applies to any kind of practice, whether everyday games like chess, or highly complex systems of practices such as we find in the sciences. Indeed, for the sciences, there seems to be much more at stake in the kind of responsibility required, for the sciences do not just purport to provide intelligible interpretations of the world; rather, our scientific understanding must be accountable to how the world is. The existential nature of the responsibility that Haugeland thinks underpins scientific understanding can be seen in his distinction between the three levels of self-criticism that were discussed in Chapter Five. At the first and second levels, the responsibility involved pertains, respectively, to (a) enforcing the norms endorsed in a scientific practice or by a community, and (b) to modifying the norms and standards of a practice in order to improve the likelihood of its success. The third and final level of self-criticism involves the willingness to (c) put an entire scientific practice into question. That is, Haugeland thinks that scientists must be prepared to abandon a set of practices or research programme if its norms and standards prevent it from being accountable to the world. This level of self-criticism involves a distinctly existential commitment, in that

> […] it is an honest commitment – in the sense of resolve or dedication – to making something work, on pain of having to give the whole thing up. Such honest commitment is "double-edged" – it cuts both ways – in that, first, it requires honest and dedicated effort to making it work, and yet, second, it also requires the honest courage, eventually, to admit that it cannot be made to work – if it cannot – and then to quit. (Haugeland 2013: 274)

The reason Haugeland thinks that the third level involves a distinct existential dimension is because it threatens the very being of the scientist as scientist: 'Who, after all, are you, professionally, if your professional specialty dies?' (ibid: 271).[2] In this respect, we can see -

[2] Although these issues might sound at some remove from the history and philosophy of science, Stephen Gaukroger (2006, 2012) has demonstrated that at key points in the history of science,

as Heidegger did - that my understanding of things in our norm-governed practices and my self-understanding are entwined. As we saw in Chapter One, how I understand things in the world is ultimately tied to my commitment to being 'for-the-sake-of' something, paradigmatically in the form of some practical identity, such as being a writer, or in this case, being a scientist. For Heidegger, in being for-the-sake-of-something, I take responsibility not only for acting in accord with the norms and standards that define a particular practical identity and its relation to various social practices; I also take a responsibility to see those norms as norms, to treat them as something that can become an issue for me, and to respond *authentically* to apparent breakdowns of them.

debates within sciences concerning methods, results, and theories were intimately bound up with larger debates about what it means to be a scientist. A useful way to develop the claims set out in this conclusion therefore would be to discuss Gaukroger's thesis in the light of these proposals.

Bibliography

Achinstein, P. 1983. *The Nature of Explanation*. New York: Oxford University Press.

Ackermann, R. 1989. The New Experimentalism. *The British Journal for the Philosophy of Science* 40 (2): 185-190.

Andersen, H. 2016. Collaboration, interdisciplinarity, and the epistemology of contemporary science. *Studies in History and Philosophy of Science* 56: 1-10.

Ankeny, R. A. 2009. Model organisms as fictions. In M. Suàrez (ed.), *Fictions in science*. New York: Routledge. 193-204.

Ankeny, R. A., Chang, H., Boumans, M., and Boon, M. 2011. Introduction: philosophy of science in practice. *European Journal of Philosophy of Science* 1: 303-307.

Ankeny, R. A., and Leonelli, S. 2011. What's so special about model organisms? *Studies in History and Philosophy of Science* 42: 313-323.

Ankeny, R. A., and Leonelli, S. 2016. Repertoires: A post-Kuhnian perspective on scientific change and collaborative research. *Studies in History and Philosophy of Science* 60: 18-28.

Atkins, P., and De Paula, J. 2014. *Atkins' Physical Chemistry*. Oxford: Oxford University Press.

Aydinonat, N. E. 2007. Models, conjectures and exploration: an analysis of Schelling's checkerboard model of residential segregation. *Journal of Economic Methodology* 14(4): 429-454.

Bailer-Jones, D. M. 2009. *Scientific Models in Philosophy of Science*. Pittsburgh, PA: University of Pittsburgh Press.

Barnes, B. 2001. Practice as Collective Action. In T.R. Schatzki, K. Knorr-Cetina, and E. von Savigny. *The Practice Turn in Contemporary Theory*. London: Routledge. 25-36.

Baumberger, C. 2011. Types of Understanding: Their Nature and Their Relation to Knowledge. *Conceptus* 40: 67-88.

Baumberger, C., Beisbart, C., Brun, G. 2016. What is Understanding? An Overview of Recent Debates in Epistemology and Philosophy of Science. In S. R. Grimm, C. Baumberger, and S. Ammon (eds.) *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*. Oxford: Routledge. 1-34.

Bechtel, W. and Richardson, R. C. 1993. *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton: Princeton University Press.

Bird, A. 2010. Social Knowing: The Social Sense of 'Scientific Knowledge'. *Philosophical Perspectives* 24: 23-56.

Black, M. 1962. *Models and Metaphor*. Ithaca, New York: Cornell University Press.

Block, N. 1998. Semantics, conceptual role. In *Routledge Encyclopedia of Philosophy*. Taylor and Francis. <https://www.rep.routledge.com/articles/thematic/semantics-conceptual-role/v-1>. doi:10.4324/9780415249126-W037-1.

Bogen, J. and Woodward, J. 1988. Saving the phenomena. *The Philosophical Review* 97(3): 303-352.

Boghossian, P. 1989. The Rule-Following Considerations. *Mind* 98(392): 507-549.

Boghossian, P. 1993. The Normativity of Content. *Philosophical Issues* 13: 31-45.

Boon, M. and Knuuttila, T. 2008. Models as epistemic tools in engineering sciences: a pragmatic approach. In A. Meijers (ed.), *Philosophy of Technology and Engineering Sciences (Handbook of the Philosophy of Science)*, vol. 9. Amsterdam: Elsevier: 687-720.

Boumans, M. 1999. Built-in Justification. In M. Morgan and M. Morrison (eds.), *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press. 66-96.

Bowler, P. J. 2003. *Evolution: The History of an Idea*. Berkeley: University of California Press.

Brandom, R. 1983. Heidegger's categories in *Being and Time*. *The Monist* 66(3): 387-409.

Brandom, R. 1994. *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Brandom, R. 2000. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Brandom, R. 2007. Inferentialism and Some of its Challenges. *Philosophy and Phenomenological Research* 74(3): 651-676.

Brandom, R. 2008. *Between Saying and Doing: Towards an Analytic Pragmatism*. Oxford: Oxford University Press.

Brigandt, I. 2010. Scientific Reasoning is Material Inference: Combining Confirmation, Discovery and Explanation. *International Studies in the Philosophy of Science* 24(1): 31-43.

Burian, R. M. 1997. Exploratory Experimentation and the Role of Histochemical Techniques in the Work of Jean Brachet, 1938-1952. *History and Philosophy of the Life Sciences* 19(1): 27-45.

Cartwright, N. 1983. *How the Laws of Physics Lie*. Oxford: Oxford University Press.

Cartwright, N. 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Cartwright, N. and Bhakthavatsalam, S. 2017. What's so special about empirical adequacy? *European Journal for Philosophy of Science* 7: 445-465.

Cash, M. 2008. Thoughts and oughts. *Philosophical Explorations* 11(2): 93-119.

Chang, H. 2004. *Inventing Temperature*. Oxford: Oxford University Press.

Chang, H. 2012. *Is Water H_2O? Evidence, Realism and Pluralism*. Dordrecht: Springer.

Clark, A. 2011. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.

Contessa, G. 2007. Scientific Representation, Interpretation and Surrogative Reasoning. *Philosophy of Science* 74(1): 48-68.

Craver, C. F. 2007. *Explaining the Brain*. Oxford: Oxford University Press.

Craver, C. F. and Darden, L. 2013. *In Search of Mechanisms: Discoveries across the Life Sciences*. Chicago: University of Chicago Press.

Crowell, S. 2002. Is There a Phenomenological Research Program? *Synthese* 131(3): 419-444.

Crowell, S. 2013. *Normativity and Phenomenology in Husserl and Heidegger*. Cambridge: Cambridge University Press.

Cushing, J. 1994. *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony*. Chicago: University of Chicago Press.

Da Costa, N. and French, S. Models, Theories, and Structures: Thirty Years On. In D. Howard (ed.) *PSA 1998, Philosophy of Science Supplement* 67: S116-S127.

Davidson, D. 1986. A Nice Derangement of Epitaphs. In E. Lepore (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell.

Daston, L. 1992. Objectivity and the Escape from Perspective. *Social Studies of Science* 22: 597-618.

Daston, L., and Galison, P. 1992. The Image of Objectivity. *Representations* 40: 81-128.

Daston, L., and Galison, P. 2010. *Objectivity*. New York: Zone Books.

Dear, P. 2006. *The Intelligibility of Nature: How Science Makes Sense of the World.* Chicago: University of Chicago Press.

Dellsen, F. 2016. Understanding without Justification or Belief. *Ratio* 30(3): 239-254.
de Donato Rodriguez, X., and Zamora Bonilla, J. 2009a. Credibility, Idealization, and Model Building: An Inferential Approach. *Erkenntnis*, 70(1): 101–18.

de Donato Rodriguez, X., and Zamora Bonilla, J. 2009b. Explanation and Modelization in a Comprehensive Inferential Account. In *EPSA09: 2nd Conference of the European Philosophy of Science Association* (Amsterdam, 21-24 October 2009). 33-42.

De Regt, H. W. 2009a. Understanding and Scientific Explanation. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press. 21-42.

De Regt, H. W. 2009b. The Epistemic Value of Understanding. *Philosophy of Science* 76(5): 585-597.

De Regt, H. W. 2013. Understanding and explanation: Living apart together? *Studies in History and Philosophy of Science* 44: 505-509.

De Regt, H. W. 2015. Scientific understanding: truth or dare? *Synthese* 192: 3781-3797.

De Regt, H. W. and Dieks, D. 2005. A Contextual Approach to Scientific Understanding. *Synthese* 144: 137-170.

De Regt, H. W., Leonelli, S., and Eigner, K. 2009. Focusing on scientific understanding. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press. 1-20.

Dreyfus, H. 1991. *Being-in-the-world: Commentary on Heidegger's* Being and Time, *Division 1*. Cambridge, MA: The MIT Press.

Dummett, M. 1973. *Frege: Philosophy of Language*. Cambridge, MA: Harvard University Press.

Elgin, C. 1993. Understanding: Art and Science. *Synthese* 95(1): 13-28.

Elgin, C. 2004. True Enough. *Philosophical Issues* 14: 113-131.

Elgin, C. 2007. Understanding and the Facts. *Philosophical Studies* 132(1): 33-42.

Elgin, C. 2009. Exemplification, Idealization, and Scientific Understanding. In Mauricio Suárez (ed.), *Fictions in Science: Philosophical Essays on Modelling and Idealization*. New York: Routledge. 77-90.

Faye, J. 2007. The Pragmatic-Rhetorical Theory of Explanation. In J. Persson and P. Ylikoski (eds.), *Rethinking Explanation*. Dordrecht: Springer.

Faye, J. 2014. *The Nature of Scientific Thinking: On Interpretation, Explanation, and Understanding*. London: Palgrave Macmillan.

Feest, U. 2010. Concepts as Tools in the Experimental Generation of Knowledge in Cognitive Neuropsychology. *Spontaneous Generations: A Journal for the History and Philosophy of Science* 4(1): 173-190.

Feest, U. 2011. What exactly is stabilized when phenomena are stabilized? *Synthese* 182: 57-71.

Feest, U. 2016. The experimenters' regress reconsidered: Replication, tacit knowledge, and the dynamics of knowledge generation. *Studies in History and Philosophy of Science* 58: 34-45.

Feest, U. and Steinle, F. 2012. *Scientific Concepts and Investigative Practice*. Berlin: Walter de Gruyter.

Fennell, J. 2013. The Meaning of 'Meaning is Normative'. *Philosophical Investigations* 36(1): 56-78.

Fernandez, A.V. 2017. The subject matter of phenomenological research: existentials, modes, and prejudices. *Synthese* 194 (9): 3543-3562.

Fodor, J. and Lepore, E. 2001. Brandom's Burdens: Compositionality and Inferentialism. *Philosophy and Phenomenological Research* 63(2): 465-483.

Franklin, A. 1989. *The Neglect of Experiment*. Cambridge: Cambridge University Press.

French, S. and Ladyman, J. 1999. Reinflating the Semantic Approach. *International Studies in the Philosophy of Science*. 13(2): 103-121.

Friedman, M. 1974. Explanation and Scientific Understanding. *The Journal of Philosophy* 71: 5-19.

Friedman, M. 2000. *A Parting of the Ways: Carnap, Cassirer, and Heidegger*. Pittsburgh: Carus Publishing Company.

Frigg, R. 2010. Models and Fiction. *Synthese* 172: 251-268.

Frigg, R. 2012. *Models in Science*. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. URL = <https://plato.stanford.edu/archives/sum2018/entries/models-science/>.

Galison, P. 1987. *How Experiments End*. Chicago: University of Chicago Press.

Galison, P. 1997. *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press.

Gaukroger, S. 2006. *The Emergence of a Scientific Culture: Science and the Shaping of Modernity 1210-1685*. Oxford: Oxford University Press.

Gaukroger, S. 2012. *The Collapse of Mechanism and the Rise of Sensibility: Science and the Shaping of Modernity 1680-1760*. Oxford: Oxford University Press.

Gelfert, A. 2016. *How to Do Science with Models: A Philosophical Primer*. Dordrecht: Springer.

Gibbard, A. 2012. *Meaning and Normativity*. Oxford: Oxford University Press.

Giere, R. N. 1988. *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press.

Giere, R. N. 2006. *Scientific Perspectivism*. Chicago: Chicago University Press.

Gijsbers, V. 2013. Understanding, Explanation, and Unification. *Studies in History of Philosophy of Science* 44: 516-522.

Ginsborg, H. 2012. The Normativity of Meaning. *Proceedings of the Aristotelian Society* (Supplementary Volume) 86: 127-146.

Glazebrook, P. 2000. *Heidegger's Philosophy of Science*. New York: Fordham University Press.

Glazebrook, P. 2012. *Heidegger on Science*. New York: State University of New York Press.

Glüer and Wilkforss. 2015. The Normativity of Meaning and Content. In. E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. URL = <https://plato.stanford.edu/archives/spr2018/entries/meaning-normativity/>.

Godfrey-Smith, P. 2003. *Theory and Reality: An Introduction to the Philosophy of Science*. Chicago:
University of Chicago Press.

Godfrey-Smith, P. 2006. The strategy of model-based science. *Biology and Philosophy* 21: 725-740.

Goodman, N. 1954. *Facts, Fiction and Forecast*. University of London: Athlone Press.

Grimm, S. 2006. Is Understanding a Species of Knowledge? *The British Journal for the Philosophy of Science* 57(3): 515-535.

Grimm, S. 2010. The goal of explanation. *Studies in History and Philosophy of Science* 41(4): 337-344.

Grimm, S. 2011. Understanding. In Sven Bernecker and Duncan Pritchard (eds.), *The Routledge Companion to Epistemology*. New York: Routledge. 84-94.

Grimm, S. 2012. The Value of Understanding. *Philosophy Compass* 7(2): 103-117.

Hacking, I. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.

Hacking, I. 2002. *Historical Ontology*. Cambridge, MA: Harvard University Press.

Hart, H.L.A. 1961. *The Concept of Law*. Oxford: The Clarendon Press.

Hattiangadi, A. 2006. Is meaning normative? *Mind & Language* 21: 220-240.

Haugeland, J. 1998. *Having Thought*. Cambridge: Harvard University Press.

Haugeland, J. 2007. Letting Be. In S. Crowell and J. Malpas (eds.), *Transcendental Heidegger*. Stanford, CA: Stanford University Press. 93-103.

Haugeland, J. 2013. *Dasein Disclosed: John Haugeland's Heidegger*. Cambridge, MA: Harvard University Press.

Haugeland, J. 2017. Two Dogmas of Rationalism. In Z. Adams and J. Browning (eds.), *Giving a Damn: Essays in Dialogue with John Haugeland*. Cambridge, MA: The MIT Press. 293-310.

Heidgger, M. 1967. *What is a Thing?* (trans. W. Barton and V. Deutsch). Chicago: Regnery.

Heidegger, M. 2008 [1962]. *Being and Time*. (trans. J. Macquarrie and E. Robinson). New York: Harper & Row, Publishers, Inc.

Heidegger, M. 2011. *Basic Writings*. Oxford: Routledge.

Hempel, C. 1965. *Aspects of Scientific Explanation and other essays in the philosophy of science*. New York: Free Press.

Hills, A. 2015. Understanding Why. *Noûs*. 50(4): 661-688.

Hutchins, E. 1995. *Cognition in the Wild*. Cambridge, MA: The MIT Press.

Hughes, R.I.G. 1997. Models and Representation. *Philosophy of Science* (suppl.) 64: S325-S336.

Humphreys, P. 2004. *Extending Ourselves: Computational Science, Empiricism and Scientific Method*. Oxford: Oxford University Press.

Keller, E. F. 2000. Models of and Models for: Theory and Practice in Contemporary Biology. *Philosophy of Science* (suppl.) 67: S72-S86.

Kellert, S. H., Longino, H., and Waters, C. K. 2006. Introduction: The pluralist stance. In S. H. Kellert, H. Longino, and C. K. Waters (eds.), *Scientific pluralism. Minnesota Studies in Philosophy of Science: Vol XIX*. Minneapolis: University of Minnesota Press. vii-xxix.

Kelp, C. 2015. Understanding phenomena. *Synthese* 192: 3799-3816.

Khalifa, K. 2012. Inaugurating understanding or repackaging explanation. *Philosophy of Science* 79: 15-37.

Khalifa, K. 2013. The role of explanation in understanding. *The British Journal for the Philosophy of Science* 64: 161-187.

Kitcher, P. 1981. Explanatory unification. *Philosophy of Science* 48: 507-531.

Kitcher, P. 1989. Explanatory unification and the causal structure of the world. In P. Kitcher and W. Salmon (eds.), *Scientific Explanation*. Minneapolis: University of Minnesota Press.

Knuuttila, T. 2011. Modelling and representing: An artefactual approach to model-based representation. *Studies in the History and Philosophy of Science* 42: 262-271.

Knuuttila, T. and Merz, M. 2009. Understanding by Modeling: An Objectual Approach. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press. 146-168.

Kohler, R. 1994. *Lords of the Fly:* Drosophila *Genetics and the Experimental Life*. Chicago: University of Chicago Press.

Kripke, S. 1981. *Naming and Necessity*. Oxford: Wiley-Blackwell.

Kuhn, T. 1979. *The Essential Tension*. Chicago: University of Chicago Press.

Kuhn, T. 1996. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kukla, R., and Lance, M. *'Yo!' and 'Lo!': The Pragmatic Topography of the Space of Reasons*. Cambridge, MA: Harvard University Press.

Kuorikoski, J. 2011. Simulation and the sense of understanding. In P. Humphreys and C. Imbert (eds.), *Models, Simulations, and Representations*.

Kuorikoski, J., and Lehtinen, A. 2009. Incredible Worlds, Credible Results. *Erkenntnis* 70: 119-131.

Kuorikoski, J., and Ylikoski, P. 2015. External representations and scientific understanding. *Synthese* 192: 3817-3837.

Kusch, M. 2002. *Knowledge by Agreement: The Programme of Communitarian Epistemology*. Oxford: Oxford University Press.

Kvanvig, J. 2003. *The Value of Knowledge and the Pursuit of Understanding*. New York: Cambridge University Press.

Kvanvig, J. 2009. The Value of Understanding. In A. Haddock, A. Millar, and D. Pritchard (eds.), *Epistemic Value*. Oxford: Oxford University Press. 95-111.

Lance, M. 2008. Placing in a Space of Norms: Neo-Sellarsian Philosophy in the Twenty-first Century. In C. Misak (ed.), *The Oxford Handbook of American Philosophy*. Oxford: Oxford University Press. 403-429.

Lance, M., and O'Leary-Hawthorne, J. 1997. *The Grammar of Meaning: Normativity and Semantic Discourse*. Cambridge: Cambridge University Press.

Latour, B., and Woolgar, S. 1979. *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.

Lenhard, J. 2009. The Great Deluge: Simulation Modelling and Scientific Understanding. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press. 169-188.

Leonelli, S. 2009a. Understanding in Biology: The Impure Nature of Biological Knowledge. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press. 189-209.

Leonelli, S. 2009b. On the Locality of Data and Claims about Phenomena. *Philosophy of Science*. 76(5): 737-749.

Leonelli, S. 2013. Integrating data to acquire new knowledge: Three modes of integration in plant science. *Studies in History and Philosophy of Biological and Biomedical Sciences*. 44: 503-514.

Leonelli, S. 2014. Data Interpretation in the Digital Age. *Perspectives on Science* 22(3): 397-417.

Lipton, P. 1991. *Inference to the Best Explanation*. Oxford: Routledge.

Lipton, P. 2009. Understanding without Explanation. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press. 43-63.

Longino, H. 1992. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton, NJ: Princeton University Press.

Longino, H. 2001. *The Fate of Knowledge*. Princeton, NJ: Princeton University Press.

Love, A. C. 2015. Collaborative explanation, explanatory roles, and scientific explaining in practice . *Studies in History and Philosophy of Science* 52: 88-94.

Machamer, P., Darden, L., and Craver, C. 2000. Thinking about Mechanisms. *Philosophy of Science* 67: 1-25.

MacIntyre, A. 1980. Epistemological crises, dramatic narrative, and the philosophy of science. In G. Gutting (ed.) *Paradigms and Revolutions*. Notre Dame, IN: University of Notre Dame Press. 54-74.

Mahan, B. H. 1975. *University Chemistry*. Reading, MA: Addison-Wesley Publishing Company.

Maher, C. 2012. *The Pittsburgh School of Philosophy: Sellars, McDowell, Brandom*. Oxford: Routledge.

Mäki, U. 2005. Models are experiments, experiments are models. *Journal of Economic Methodology* 12(2): 303-315.

McDaniel, K. 2013. Heidegger's Metaphysics of Material Beings. *Philosophy and Phenomenological Research* 87(2): 332-357.

McDowell, J. 1984. Wittgenstein on Following a Rule. *Synthese* 58(3): 325-363.

McDowell, J. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.

McDowell, J. 2009. *The Engaged Intellect: Philosophical Essays*. Cambridge, MA: Harvard University Press.

McGinn, M. 1997. *The Routledge Guidebook to Wittgenstein's* Philosophical Investigations. Oxford: Routledge.

McManus, D. 2012. *Heidegger and the Measure of Truth*. Oxford: Oxford University Press.

McManus, D. 2015. *Heidegger, Authenticity and the Self: Themes from Division Two of* Being and Time. Oxford: Routledge.

Mizrahi, M. 2012. Idealizations and scientific understanding. *Philosophical Studies* 160: 237-252.

Moore, A.W. 2012. *The Evolution of Metaphysics: Making Sense of Things*. Cambridge: Cambridge University Press.

Morgan, M. 2003. Experiments without Material Intervention: Model Experiments, Virtual Experiments and Virtually Experiments. In H. Radder (ed.) *The Philosophy of Scientific Experimentation*. Pittsburgh: University of Pittsburgh Press. 217-235.

Morgan, M., and Boumans, M. 2004. Secrets Hidden by Two-Dimensionality: The Economy as a Hydraulic Machine. In S. de Chadarevian and N. Hopwood (eds.), *Models: The Third Dimension of Science*. Stanford: Stanford University Press. 369-401.

Morgan, M., and Morrison, M. 1999. *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.

Morrison, M. 2011. One phenomenon, many models: Inconsistency and complementarity. *Studies in the History and Philosophy of Science* 42: 342-351.

Myers, N. 2008. Molecular Embodiments and the Body-work of Modeling in Protein Crystallography. *Social Studies of Science* 38(2): 163-199

Nersessian, N. J. 2006. Model-Based Reasoning in Distributed Cognition. *Philosophy of Science* 73(4): 699-709.

Nersessian, N. J. 2008. *Creating Scientific Concepts*. Cambridge: MIT Press.

Newman, M. 2012. An Inferential Model of Scientific Understanding. *International Studies in the Philosophy of Science* 26: 1-26.

Newman, M. 2013. EMU and Inference: what the explanatory model of scientific understanding ignores. *European Journal for Philosophy of Science* 4: 55-74.

Noë, A. 2012. *Varieties of Presence*. Cambridge, MA: Harvard University Press.

Nounou, A., and Psillos, S. 2012. Book review of *Scientific understanding: Philosophical perspectives*. In H. W. de Regt, S. Leonelli, K. Eigner (Eds.), *Studies in history and philosophy of modern physics*, 43: 72–74.

Norton, J. 2003. A Material Theory of Induction. *Philosophy of Science* 70(4): 647-670.

Norton, J. 2010. There Are No Universal Rules of Induction. *Philosophy of Science* 77(5): 765-777.

O'Shea, J. 2010. Normativity and Naturalism in Sellars' 'Janus-Faced' Space of Reasons. *International Studies of Philosophical Studies* 18(3): 459-471.

Okrent, M. 1988. *Heidegger's Pragmatism: Understanding, Being, and the Critique of Metaphysics*. Ithaca, New York: Cornell University Press.

Peregrin, J. 2012. Inferentialism and the Normativity of Meaning. *Philosophia*, 40: 75–97.

Peregrin, J. 2014. *Inferentialism: Why Rules Matter*. London: Palgrave Macmillan.

Peschard, I. 2007. Participation of the Public in Science: Towards a New Kind of Scientific Practice. *Human Affairs* 17: 138-153.

Pickering, A. 1995. *The Mangle of Practice: Time, Agency and Science*. Chicago: University of Chicago Press.

Polanyi, M. 1958. *Personal Knowledge: Towards a post-critical philosophy*. London: Routledge & Kegan Paul.

Polt, R. 1999. *Heidegger: An Introduction*. London: UCL Press.

Popper, K. 2002. *The Logic of Scientific Discovery*. Oxford: Routledge.

Radder, H. 2003. *The Philosophy of Scientific Experimentation*. Pittsburgh: University of Pittsburgh Press.

Rehg, W. 2009. *Cogent Science in Context: The Science Wars, Argumentation Theory, and Habermas*. Cambridge, MA: The MIT Press.

Reiss, J. 2012. Causation in the Sciences: An Inferentialist Account. *Studies in History and Philosophy of Science Part C* 43(4): 769-777.

Rheinberger, H. 1997. *Towards a History of Epistemic Things*. Stanford: Stanford University Press.

Rietveld, E. 2008. Situated Normativity: The Normative Aspect of Embodied Cognition in Unreflective Action. *Mind* 117(468): 973-1001.

Riggs, W. 2003. Understanding 'Virtue' and the Virtue of Understanding. In M. DePaul and L. Zagzebski (eds.), *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Clarendon Press. 203-226.

Rorty, R. 1979. *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press.

Rouse, J. 1981. Kuhn, Heidegger and Scientific Realism. *Man and World* 14: 269-290.

Rouse, J. 1987. *Knowledge and Power: toward a political philosophy of science*. Ithaca, New York: Cornell University Press.

Rouse, J. 2002. *How Scientific Practices Matter*. Chicago: University of Chicago Press.

Rouse, J. 2003. Kuhn's Philosophy of Scientific Practice. In T. Nickles (ed.) *Thomas Kuhn*. Cambridge: Cambridge University Press.

Rouse, J. 2005. Heidegger's Philosophy of Science. In H. Dreyfus and M. Wrathall (eds.), *A Companion to Heidegger*. Oxford: Blackwell Publishing Ltd. 173-190.

Rouse, J. 2007. Social Practices and Normativity. *Philosophy of the Social Sciences* 37(1): 46-56.

Rouse, J. 2009. Laboratory Fictions. In M. Suarez (ed.), *Fictions in Science: Philosophical Essays on Modelling and Idealization*. New York: Routledge. 37-55.

Rouse, J. 2011. Articulating the world: Experimental systems and conceptual understanding. *International Studies in Philosophy of Science* 25: 243-254.

Rouse, J. 2013. Recovering Thomas Kuhn. *Topoi* 32: 59-64.

Rouse, J. 2015a. *Articulating the World: Conceptual Understanding and the Scientific Image*. Chicago: University of Chicago Press.

Rouse, J. 2015b. Scientific Law, Natural Necessity, and Heideggerian Ontology. *Journal of Dialectics of Nature* 37(5): 89-101.

Rouse, J. 2016. Towards a New Naturalism: Niche Construction, Conceptual Normativity, and Scientific Practice. In M. Risjord (ed.), *Normativity and Naturalism in the Philosophy of the Social Sciences*. Oxford: Routledge. 28-42.

Rouse, J. 2017. Normativity. In Julian Kiverstein (ed.) *The Routledge Handbook of Philosophy of the Social Mind*. Oxford: Routledge. 545-562.

Ryle, G. 1946. Knowing How and Knowing That. *Proceedings of the Aristotelian Society* 46: 1-16.

Salmon, W. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

Schatzki, T., Knorr Cetina, K., and von Savigny, E. 2001. *The Practice Turn in Contemporary Theory*. New York: Routledge.

Schear, J. K. 2013. Understanding as Finite Ability. In D. Egan, S. Reynolds, A.J. Wendland (eds.) *Wittgenstein and Heidegger*. 163-178.

Schelling, T. 1978. *Micromotives and macrobehavior*. London & New York, NY: W. W. Norton.

Scriven, M. 1962. Explanation, Predictions and Laws. In *Minnesota Studies in the Philosophy of Science* III.

Sellars, W. 1948. Concepts as Involving Laws and Inconceivable without Them. *Philosophy of Science* 15(4): 287-315.

Sellars, W. 1953. Inference and Meaning. *Mind* 62(247): 313-338.

Sellars, W. 1954. Some Reflections on Language Games. *Philosophy of Science* 21(3): 204-228.

Sellars, W 1974. Meaning as Functional Classification. *Synthese* 27(3/4): 417-437.

Sellars, W. 1997. *Empiricism and the Philosophy of Mind*. Cambridge, MA: Harvard University Press.

Shrum, W., Genuth, J., and Chompalov, I. *Structures of Scientific Collaboration*. Cambridge, MA: The MIT Press.

Steinle, F. 1997. Entering New Fields: Exploratory Uses of Experimentation. *Philosophy of Science* (suppl.) 64: S64-S74.

Strevens, M. 2013. No Understanding without Explanation. *Studies in the History and Philosophy of Science* 44: 510-515.

Stuart, M. T. 2016. Taming theory with thought experiments: Understanding and scientific progress. *Studies in History and Philosophy of Science* 58: 24-33.

Suárez, M. 2004. An Inferential Conception of Scientific Representation. *Philosophy of Science* (suppl.) 71: S767-S779.

Suárez, M. 2015. Deflationary representation, inference, and practice. *Studies in History and Philosophy of Science* 49: 36-47.

Swoyer, C. 1991. Structural Representation and Surrogative Reasoning. *Synthese* 87: 449-508.

Taylor, C. 1985. *Philosophical Papers: Philosophy and the Human Sciences Volume 2*. Cambridge: Cambridge University Press.

Thagard, P. 1992. *Conceptual Revolutions*. Princeton, N.J: Princeton University Press.

Thompson, E. 2010. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: The Belknap Press.

Toon, A. 2012. *Models as Make-Believe: Imagination, Fiction and Scientific Representation*. London: Palgrave Macmillan.

Toon, A. 2015. Where is the understanding? *Synthese* 192: 3859-3875.

Traweek, S. 1992. *Beamtimes and Lifetimes: The World of High Energy Physics*. Cambridge, MA: Harvard University Press.

Trout, J. D. 2002. Scientific explanation and the sense of understanding. *Philosophy of Science* 69: 212-233.

Trout, J. D. 2007. The Psychology of Scientific Explanation. *Philosophy Compass* 2(3): 564-591.

Turner, S. 1994. *The Social Theory of Practices*. Chicago: University of Chicago Press.

Van Fraassen, B. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Vorms, M. 2011. Representing with imaginary models: Formats matter. *Studies in History and Philosophy of Science* 42: 287-295.

Wanderer, J. 2008. *Robert Brandom*. Oxford: Routledge.

Waters, C. K. 2004. What was classical genetics? *Studies in History and Philosophy of Science* 35: 783-809.

Weber, M. 2005. *Philosophy of Experimental Biology*. Cambridge: Cambridge University Press.

Weisberg, M. 2007. Who is a Modeler? *The British Journal for the Philosophy of Science* 58(2): 207-233.

Weisberg, M. 2013. *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

Wheeler, M. 2017. Martin Heidegger. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/fall2017/entries/heidegger/>.

Whiting, D. 2009. Is Meaning Fraught with Ought? *Pacific Philosophical Quarterly* 90: 535-555.

Wilkenfeld, D. A. 2013. Understanding as Representation Manipulability. *Synthese* 190: 997-1016.

Wilkenfeld, D. A. 2017. MUDdy understanding, *Synthese* 194: 1273-1293.

Wilkenfeld, D. A., and Hellmann, J.K. 2014. Understanding beyond grasping propositions: A discussion of chess and fish, *Studies in History and Philosophy of Science* 48: 46-51.

Williams, M. 2001. *Problems of Knowledge: A Critical Introduction to Epistemology*. Oxford: Oxford University Press.

Williams, M. 2008. Responsibility and Reliability. *Philosophical Papers* 37(1): 1-26.

Wilson, M. 2006. *Wandering Significance: An Essay on Conceptual Behaviour*. Oxford: Oxford University Press.

Winsberg, E. 2006. Models of success versus the success of models: reliability without truth. *Synthese* 152: 1-19.

Winsberg, E., Huebner, B., and Kukla, R. 2014. Accountability and values in radically collaborative research. *Studies in History and Philosophy of Science* 46: 16-23.

Withy, K. 2011. Situation and Limitation: Making Sense of Heidegger on Thrownness. *European Journal of Philosophy* 22(1): 61-81.

Withy, K. 2017. Haugeland's Heidegger and the Metaphysics of Normativity. *European Journal of Philosophy* 25(2): 463-484.

Wittgenstein, L. 2009. *Philosophical Investigations*. (trans. G.E.M. Anscombe, P.M.S. Hacker and J. Schulte). Oxford: Blackwell Publishing Ltd.

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press

Woody, A. 2004. Telltale Signs: What Common Explanatory Strategies in Chemistry Reveal About Explanation Itself. *Foundations of Chemistry* 6: 13-43.

Woody, A. 2013. How is the Ideal Gas Law Explanatory? *Science and Education* 22: 1563-1580.

Woody, A. 2014. Chemistry's Periodic Law: Rethinking Representation and Explanation After the Turn to Practice. In L. Soler, S. Zwart, M. Lynch, V. Israel-Jost, *Science After the Practice Turn in the Philosophy, History, and Social Studies of Science*. Oxford: Routledge. 123-150.

Woody, A. 2015. Re-orienting discussions of scientific explanation: A functional perspective. *Studies in History and Philosophy of Science* 52: 79-87.

Wray, K. B. 2001. Collective Belief and Acceptance. *Synthese* 129(3): 319-333.

Wray, K. B. 2002. The Epistemic Significance of Collaborative Research. *Philosophy of Science*. 69(1): 150-168.

Ylikoski, P. 2009. The Illusion of Depth of Understanding. In H. W. de Regt, S. Leonelli, and K. Eigner (eds.), *Scientific Understanding: Philosophical Perspectives*. Pittsburgh, PA: University of Pittsburgh Press. 100-119.

Ylikoski, P., and Aydinonat, N. E. 2014. Understanding with theoretical models. *Journal of Economic Methodology* 21(1): 19-36.

Zagzebski, L. 2001. Recovering Understanding. In Matthias Steup (ed.), *Knowledge, Truth and Duty: Essays on Epistemic Justification, Responsibility and Virtue*. New York: Oxford University Press: 235-256.